

A Brownian particle in a microscopic periodic potential

Jeremy Clark^{*} and Loïc Dubois[†]

University of Helsinki, Department of Mathematics
Helsinki 00014, Finland

October 5, 2011

Abstract

We study a model for a massive test particle in a microscopic periodic potential and interacting with a reservoir of light particles. In the regime considered, the fluctuations in the test particle's momentum due to collisions typically outweigh the shifts in momentum due to the periodic force, and so, the force is effectively a perturbative contribution. The mathematical starting point is an idealized reduced dynamics for the test particle given by a linear Boltzmann equation. In the limit that the mass ratio of a single reservoir particle to the test particle tends to zero, we show that under the standard normalizations for the test particle variables, there is convergence to the Ornstein-Uhlenbeck process. Our analysis is primarily directed towards bounding the perturbative effect of the periodic potential on the particle's momentum.

1 Introduction

The Ornstein-Uhlenbeck process offers a homogenized picture for the motion of a massive particle interacting with a gas of lightweight particles at fixed temperature [33]. In this description, the spatial degrees of freedom are driven ballistically by momentum variables which are themselves governed by a diffusion equation that includes a drift term corresponding to the drag felt by the massive particle as it accumulates speed and has more frequent collisions with the gas. Under diffusive rescaling, the spatial variables converge in law to a Brownian motion. This result follows by an elementary analysis of the closed formulas available for the Ornstein-Uhlenbeck process [26]. The Brownian motion description for the test particle transport is effectively “more macroscopic” than the Ornstein-Uhlenbeck model, since the fluctuations in the particle's momentum are integrated into infinitesimal spatial “jumps” for the Brownian particle.

In the other direction, we may consider derivations of the Ornstein-Uhlenbeck process from models which are “more microscopic”. These relatively microscopic descriptions may merely be more complicated stochastic models for the test particle (such as a Boltzmann or linear Boltzmann equation), or, more fundamentally, a reduced dynamics for the test particle starting from a full microscopic model which includes the evolution of the degrees of freedom for the

^{*}jclark@mappi.helsinki.fi

[†]ldubois@mappi.helsinki.fi

gas. The stochastic model in the former case should be regarded as an intermediary picture between the Ornstein-Uhlenbeck and the Hamiltonian dynamics arising in some limit (see [30] for a discussion of the low density limit). In the Boltzmann models, the test particle undergoes a Markovian dynamics, whereas for the Hamiltonian model including the gas, the randomness is only in the initial configuration, and the resulting dynamics for the test particle given by integrating out the gas is non-Markovian. In the other direction, the contrast between the Ornstein-Uhlenbeck and the Boltzmann-type dynamics is that the momentum in the Boltzmann case makes discrete jumps (which are individually small in the Brownian limit) corresponding to collisions with gas particles rather than evolving with continuous trajectories according to a Langevine equation as in the Ornstein-Uhlenbeck case. We refer to the book [26] for a discussion of these various levels of description for a Brownian particle.

Rigorous mathematical derivations of the Ornstein-Uhlenbeck process were achieved in [15, 3] from stochastic models giving an effective description of the test particle receiving collisions from particles in a background gas. For models which begin with a full mechanical Hamiltonian model including the test particle and the gas, derivations of the Ornstein-Uhlenbeck process from the reduced dynamics of the test particle were obtained in [16, 10, 31].

In this article, we consider the Brownian regime for a stochastic model in which a one-dimensional test particle makes jumps in momentum (interpreted as collisions with a background gas) and is acted upon by a force from an external, spatially periodic potential field. Due to the presence of the field, the momentum process is no longer Markovian, since it drifts at a rate depending on the particle's position. The momentum of the particle has two contributions: the total displacement in momentum due to the field given by a time integral of the force and the sum of the momentum jumps from collisions. Due to the specific scaling regime considered (which includes the period length of the potential), the force field typically makes a smaller-scale contribution to the test particle's momentum than the fluctuations in momentum due to the jumps ("collisions with the gas"). The vanishing of the force contribution is an averaged effect driven by the frequent rate at which the test particle is typically passing through the period cells of the potential field. The Brownian limit of the model to first-order thus yields the same Ornstein-Uhlenbeck process as if the force were set to zero. Our analysis is focused on obtaining a sharp upper bound for the influence of the external potential on the momentum of the particle, and our techniques improve those applied to a related model in [9]. Ultimately, the main contributions to the total drift in momentum due to the forcing are made during "rare" time periods at which the test particle's momentum returns to "small" values. The results of this article are extended in [7] to prove that the integral of the force (or net displacement in momentum due to the potential) converges in law to a fractional diffusion whose rate depends on the amount time that the limiting Ornstein-Uhlenbeck process spends at zero momentum (i.e. the local time at zero).

Our model is a linear Boltzmann dynamics for a one-dimensional particle making elastic collisions with the gas and including a spatially periodic potential. The jump rate kernel is the one-dimensional case of the formula appearing in [30, Ch. 8.6], which corresponds to a hard-rod interaction between the test particle and a single reservoir particle. However, since the model is one-dimensional, it cannot be derived from a mechanical microscopic dynamics in the Boltzmann-Grad limit. We thus regard our model as phenomenological, and we argue that the resulting behavior that we find is qualitatively the same as what should be expected in an analogous three-dimensional model for a Brownian particle in a one-dimensional periodic potential.

We think of our model as corresponding to an experimental situation for a large atom or

molecule in a periodic standing-wave light field (optical lattice) and interacting with a dilute background gas. A periodic optical force on an atom can be produced experimentally by counter-propagating lasers (see, for instance, [22] or the reviews [1, 25]). A classical treatment of the atom is reasonable in the regime where the potential is effectively weak, since the test particle is typically not constrained by the potential and the coherent quantum effects for the test particle will be suppressed by interactions with the gas.

1.1 Model and results

We will consider a one-dimensional particle of mass M interacting with a gas of particles with mass m for $\frac{m}{M} = \lambda \ll 1$ in the presence of a force $\frac{dV}{dx}(\frac{x}{\lambda})$ for some smooth $V : \mathbb{R} \rightarrow \mathbb{R}^+$ with period $a > 0$. We take the phase space density $\Psi_{t,\lambda}(x, p)$ at time $t > 0$ to obey a linear Boltzmann equation

$$\begin{aligned} \frac{d}{dt}\Psi_{t,\lambda}(x, p) = & -\frac{\lambda}{m}p\frac{\partial}{\partial x}\Psi_{t,\lambda}(x, p) + \frac{dV}{dx}\left(\frac{x}{\lambda}\right)\frac{\partial}{\partial p}\Psi_{t,\lambda}(x, p) \\ & + \int_{\mathbb{R}} dp' (\mathcal{J}_{\lambda}(p', p)\Psi_{t,\lambda}(x, p') - \mathcal{J}_{\lambda}(p, p')\Psi_{t,\lambda}(x, p)), \end{aligned} \quad (1.1)$$

where $\mathcal{J}_{\lambda}(p', p)$ is a kernel describing the rate of kicks in momentum $p' \rightarrow p$ for the massive particle due to collisions with reservoir particles. Since we are considering an ideal gas, the rates $\mathcal{J}_{\lambda}(p', p)$ will be determined by the interaction potential between the test particle and a reservoir particle, the temperature β^{-1} , the ratio $\lambda = \frac{m}{M}$, and the spatial density η . We will take the rates $\mathcal{J}_{\lambda}(p', p)$ to correspond to a hard-rod interaction (or alternatively “hard-point”, since the dimensionality of the objects do not appear for the one-dimensional linear Boltzmann equation) which has the form

$$\mathcal{J}_{\lambda}(p', p) = \frac{\eta(1+\lambda)}{2m} |p' - p| \frac{e^{-\frac{\beta}{2m}\left(\frac{1-\lambda}{2}p' - \frac{1+\lambda}{2}p\right)^2}}{(2\pi\frac{m}{\beta})^{\frac{1}{2}}}. \quad (1.2)$$

The jump rates \mathcal{J}_{λ} are the explicit form of those in equation (8.118) from [30] for the dimension-1 case, written in momentum variables rather than velocities.

We will denote the stochastic process whose probability density evolves according to (1.1) by (X_t, P_t) . Let us also define the process

$$D_t = \int_0^t dr \frac{dV}{dx}\left(\frac{X_r}{\lambda}\right).$$

The process D_t is the cumulative drift in the particle’s momentum due to the periodic force field, and hence the momentum at time t has the form

$$P_t = P_0 + D_t + J_t, \quad (1.3)$$

where J_t is the sum of all the momentum jumps due to collisions with the gas over the time interval $[0, t]$.

Let $(\mathbf{q}_t, \mathbf{p}_t) \in \mathbb{R}^2$ be a process satisfying the Langevine equations

$$\begin{aligned} d\mathbf{q}_t &= \frac{1}{m}\mathbf{p}_t dt, \\ d\mathbf{p}_t &= -\gamma\mathbf{p}_t dt + \left(\frac{2m\gamma}{\beta}\right)^{\frac{1}{2}}d\mathbf{B}'_t, \end{aligned} \quad (1.4)$$

where $\gamma = 8\eta\left(\frac{2}{\pi m\beta}\right)^{\frac{1}{2}}$ and \mathbf{B}'_t is a standard Brownian motion.

Since our model is rather specific, the technical assumptions for our main results are only the following list.

List 1.1.

1. The potential $V(x) = V(x + a)$ is continuously differentiable.
2. The probability measure μ on \mathbb{R}^2 for the initial location in phase space (X_0, P_0) has finite moments in momentum $\int_{\Sigma} d\mu(x, p) |p|^m < \infty$, $m \geq 1$.

The following theorem gives the main result of this article. It states that as $\lambda \rightarrow 0$ the momentum process $P_{\frac{t}{\lambda}}$ rescaled by a factor $\lambda^{\frac{1}{2}}$ converges to an Ornstein-Uhlenbeck process.

Theorem 1.2. *Assume List 1.1. In the limit $\lambda \rightarrow 0$, there is convergence in law of the process $\lambda^{\frac{1}{2}}P_{\frac{t}{\lambda}}$ to the Ornstein-Uhlenbeck process \mathbf{p}_t over the interval $[0, T]$. The convergence is with respect to the uniform metric. Moreover,*

$$\sup_{\lambda < 1} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} \left| \lambda^{\frac{1}{4}} D_{\frac{t}{\lambda}} \right| \right] < \infty.$$

1.2 Further discussion

This article concerns the dynamics of a Brownian particle which feels a light force from a one-dimensional periodic optical potential (optical lattice). There is considerable range in the period and amplitude that can be experimentally produced for an optical lattice [25], and we focus on a regime in which the potential is “microscopic”. By “microscopic”, we mean that the potential has an amplitude $\sup_{x, x'} |V(x) - V(x')|$ which is much smaller than the typical kinetic energy $\frac{M}{\beta} = \lambda^{-1} \frac{m}{\beta}$ of the test particle at equilibrium with the heat bath, and that the period a is small enough so that the typical rate at which the particle passes through the period cells $(a^2 M \beta)^{-\frac{1}{2}}$ is much faster than the rate of energy relaxation $\approx \lambda^{-1} \gamma$ for the test particle.

For our mathematical analysis, the force $F(x) = \frac{dV}{dx}\left(\frac{x}{\lambda}\right)$ is taken to have a period which scales proportionally to the mass ratio $\lambda = \frac{m}{M}$: $a\lambda$. This is not essential to these results, and only the broad features described above are critical. The same can be said about the amplitude of the potential.

Theorem 1.2 states that to first approximation under Brownian rescaling, the momentum is an Ornstein-Uhlenbeck process with no dependence on the potential. This classical treatment of the particle allows for comparisons with quantum models. A similar model for a one-dimensional quantum particle was studied in [6] for which the potential is a periodic δ -potential. In that case, the singular potential makes a first-order change to the dynamics characterized by spatial subdiffusion caused by Bragg reflections even though the periodic potential is “microscopic” in a similar sense as described above. See [4, 13, 19] for examples of experimental investigations of Bragg reflections of atoms from optical potentials. Analogous quantum models with smoother potentials will behave more like their classical counterparts.

A three-dimensional linear Boltzmann dynamics for a particle in a gas of hard spheres and under the influence of a one-dimensional optical lattice will have the same limit result (up to the constants) as in Theorem 1.2 for the degree of freedom in the direction of the potential. Although the momentum for a single spatial degree of freedom is not Markovian in the linear Boltzmann description, it becomes “more Markovian” in the Brownian limit as is seen in the

limiting three-dimensional Ornstein-Uhlenbeck process. The rates (1.2) can then be replaced by the effective rates that emerge for a single degree of freedom in the three dimensional case, and which have the same qualitative features for our purposes.

1.2.1 Features of the model

By rescaling the spatial coordinate for the particle by a factor of λ^{-1} , the master equation (1.1) becomes

$$\begin{aligned} \frac{d}{dt}\Psi_{t,\lambda}(x, p) &= \mathcal{L}_\lambda^*(\Psi_{t,\lambda})(x, p) = -\frac{p}{m}\frac{\partial}{\partial x}\Psi_{t,\lambda}(x, p) + \frac{dV}{dx}(x)\frac{\partial}{\partial p}\Psi_{t,\lambda}(x, p) \\ &\quad + \int_{\mathbb{R}} dp' (\mathcal{J}_\lambda(p', p)\Psi_{t,\lambda}(x, p') - \mathcal{J}_\lambda(p, p')\Psi_{t,\lambda}(x, p)), \end{aligned} \quad (1.5)$$

where the generator \mathcal{L}_λ^* is defined by the second equality. We thus effectively have a particle with Hamiltonian $H(x, p) = \frac{1}{2m}p^2 + V(x)$ and a λ -dependent noise. Note that under the new spatial metric, the velocity of the test particle is $\frac{p}{m}$ rather than $\frac{p}{M} = \lambda\frac{p}{m}$. For the purpose of Theorem 1.2 and this article generally, it is sufficient to consider the spatial degree of freedom to be a unit torus $\mathbb{T} = [0, 1)$ so that the total state space is $\Sigma = \mathbb{T} \times \mathbb{R}$. The equilibrium state for the dynamics on Σ is given by the Maxwell-Boltzmann distribution

$$\Psi_{\infty,\lambda}(x, p) = \frac{e^{-\beta\lambda H(x,p)}}{N(\lambda)} \quad (1.6)$$

for some normalization $N(\lambda)$.

After the spatial stretching, the drift process in momentum D_t has the form

$$D_t = \int_0^t dr g(X_r, P_r) \quad (1.7)$$

for $g : \Sigma \rightarrow \mathbb{R}^+$ given by $g(x, p) = \frac{dV}{dx}(x)$. It is thus an integral functional of an exponentially ergodic Markov process on Σ . Nonetheless, the central limit theorem for $\lambda^{\frac{1}{4}}D_{\frac{t}{\lambda}}$ stated in Theorem 1.2 does not follow from the limit theory for integral functionals of ergodic Markov processes [20], since the relaxation to the state (1.6) only occurs on the time scale $\lambda^{-1} \gg 1$. Indeed, there must be many collisions with reservoir particles before there is memory loss for the heavy particle. As explained in Sections 1.2.2-1.2.3, the analysis of $\lambda^{\frac{1}{4}}D_{\frac{t}{\lambda}}$ is more related to the limit theory for martingales whose bracket processes are additive functionals of a null-recurrent Markov process [17]. This is due to the fact that the fluctuations in D_t accumulate mainly during time intervals in which $|P_t|$ is much smaller than the typical momentum-size $(\frac{m}{\lambda\beta})^{\frac{1}{2}} \gg 1$ for the state $\Psi_{\infty,\lambda}$.

1.2.2 Rough picture of the behavior in the Brownian regime $\lambda \ll 1$

Since the equilibrium state of the dynamics is given by the Maxwell-Boltzmann distribution (1.6), the typical energy for the particle when $\lambda \ll 1$ will be on the order λ^{-1} . Moreover, the potential $V(x)$ is bounded, so most of the energy will be in the kinetic component $\frac{1}{2m}p^2$ corresponding to momenta with $|p|$ on the order of $\lambda^{-\frac{1}{2}} \gg 1$. The jump rates $\mathcal{J}_\lambda(p, p')$ for $|p| = O(\lambda^{-\frac{1}{2}})$ are approximately

$$\mathcal{J}_\lambda(p, p') = j(p - p') + \lambda p \frac{\beta}{2m} j(p - p') + O(\lambda), \quad (1.8)$$

where the idealized rates $j(p)$ have the form

$$j(p) = \frac{\eta}{2m} |p| \frac{e^{-\frac{\beta}{2m} p^2}}{(2\pi \frac{m}{\beta})^{\frac{1}{2}}}. \quad (1.9)$$

The second term on the right is $O(\lambda^{\frac{1}{2}})$ by our assumption $|p| = O(\lambda^{-\frac{1}{2}})$. The physical meaning behind the approximation (1.8) is that the gas reservoir particles are typically moving at speeds on the order $(\frac{m}{\beta})^{\frac{1}{2}}$ which is greater than the typical speed of the test particle $(\frac{M}{\beta})^{\frac{1}{2}} = \lambda^{\frac{1}{2}} (\frac{m}{\beta})^{\frac{1}{2}} \ll (\frac{m}{\beta})^{\frac{1}{2}}$. The statistics for the momentum transfers from the gas thus do not depend strongly on the momentum of the test particle, and have approximately a convolution form as in the zeroth-order term in (1.8). The zeroth-order approximation in (1.8) suggests that the collision component J_t of the momentum (1.3) is typically behaving as an unbiased random walk with increments having density $j(v)$. Based on this reasoning, $\lambda^{\frac{1}{2}} J_{\frac{t}{\lambda}}$ should converge by the central limit theorem to a Brownian motion with diffusion constant $\frac{2m\gamma}{\beta}$ as $\lambda \rightarrow 0$. However, the first-order term in (1.8) generates a drift for $\lambda^{\frac{1}{2}} J_{\frac{t}{\lambda}}$ which is retained as $\lambda \rightarrow 0$ and converges to a limit by a law of large numbers. This can be seen in the friction term appearing in the Langevine equation (1.4).

According to the heuristics above, $J_{\frac{t}{\lambda}}$ should typically be found on the scale $\lambda^{-\frac{1}{2}}$ when $\lambda \ll 1$ and $t \in [0, T]$, and we will now argue that $D_{\frac{t}{\lambda}}$ should typically be $O(\lambda^{-\frac{1}{4}})$. We can parse the integral for D_t according to the collision times t_n as

$$D_t = \int_{t_{N_t}}^t dr \frac{dV}{dx}(X_r) + \sum_{n=1}^{N_t} \int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r),$$

where $t_0 = 0$ and N_t is the number of collisions up to time t . Between collisions from the gas, the particle evolves deterministically according to the Hamiltonian $H(x, p) = \frac{1}{2}p^2 + V(x)$, and Newton's equations give

$$P_{t_n^-} - P_{t_{n-1}} = \int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r). \quad (1.10)$$

If $H(X_{t_{n-1}}, P_{t_{n-1}}) > 2 \sup_x V(x)$, then the momentum will not change signs over the interval $[t_{n-1}, t_n]$, and

$$|P_{t_n^-} - P_{t_{n-1}}| \leq \left| |P_{t_{n-1}}| - \sqrt{P_{t_{n-1}}^2 + 2V(X_{t_{n-1}}) - 2V(X_{t_n})} \right| \leq \frac{2 \sup_x V(x)}{|P_{t_{n-1}}|},$$

which follows from the conservation of energy and the quadratic formula. Thus, when $|P_{t_{n-1}}|$ is on the typical order $\propto \lambda^{-\frac{1}{2}}$, then the increment (1.10) of the momentum drift is $O(\lambda^{\frac{1}{2}}) \ll 1$.

In fact, there is another critical feature of an ergodic nature that make the contributions (1.10) to D_t even smaller when $|P_{t_{n-1}}| \gg 1$. There is an ergodicity on the spatial torus relating to the fact that when the momentum is high, then the particle revolves quickly around the torus, and its location at the time of the next collision is close to uniform. This idea can be used to show that the mean for $\int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r)$ is $O(|P_{t_{n-1}}|^{-2})$ when given the information known up to time t_{n-2} . In other words, besides the increments (1.10) just being small when

¹These velocities refer to the original length scale, before stretching by a factor of λ^{-1} .

$|P_{t_{n-1}}| \gg 1$, D_t is also behaving like a martingale since the increments are close to having mean zero and being uncorrelated. Thus, there is a central limit theorem-like cancellation among the terms. This motivates that the contribution to $D_{\frac{t}{\lambda}}$ from time intervals where $|P_r| \propto \lambda^{-\frac{1}{2}}$ is $o(\lambda^{-\frac{1}{2}})$, since for fixed small $\epsilon > 0$

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{n=1}^{\mathcal{N}_{\frac{t}{\lambda}}} \chi(|P_{t_{n-1}}| \geq \epsilon \lambda^{-\frac{1}{2}}) \int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r) \right)^2 \right] \\ &= O \left(\mathbb{E} \left[\sum_{n=1}^{\mathcal{N}_{\frac{t}{\lambda}}} \chi(|P_{t_{n-1}}| \geq \epsilon \lambda^{-\frac{1}{2}}) \left(\int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r) \right)^2 \right] \right) \leq \epsilon^{-2} \mathbb{E}[\mathcal{N}_{\frac{t}{\lambda}}] O(\lambda) = O(1), \end{aligned}$$

because the collisions occur with a frequency on the order of one per unit time $\mathbb{E}[\mathcal{N}_{\frac{t}{\lambda}}] = O(\frac{t}{\lambda})$. These contributions disappear for the normalized expression $\lambda^{\frac{1}{4}} D_{\frac{t}{\lambda}}$. (For technical reasons, our analysis of these facts is actually done with a different set of artificially introduced stopping times rather than the collision times. See Section 4.4).

The above arguments motivate that $D_{\frac{t}{\lambda}}$ spends the greater portion of the time interval $t \in [0, T]$ behaving as a constant, or, said differently, its larger fluctuations—if they occur—must typically be concentrated on a small fraction of the interval $[0, T]$. Let us consider the order of the contributions to D_t which are likely to occur for the periods of time when P_r returns to the region around the origin, that is $|P_r| = O(1)$. If P_t is behaving roughly as a random walk for $t \in [0, \frac{T}{\lambda}]$ with some very weak friction, then we expect that P_t spends on the order of $\lambda^{-\frac{1}{2}}$ time in the vicinity of the origin. If there is central limit theorem-like cancellations between the increments $\int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r)$ in those time periods, then $D_{\frac{t}{\lambda}}$ should be expected to be on the scale $\lambda^{-\frac{1}{4}}$.

1.2.3 Techniques and strategy of the proof

The main difficulty in showing that $\lambda^{\frac{1}{2}} P_{\frac{t}{\lambda}}$ converges in law to the Ornstein-Uhlenbeck process \mathbf{p}_t is to show that component $D_{\frac{t}{\lambda}}$ of the momentum is typically $o(\lambda^{-\frac{1}{2}})$ for $t \in [0, T]$. As indicated by the heuristics of Section 1.2.2, we should expect, in fact, that typically $\sup_{0 \leq t \leq T} |D_{\frac{t}{\lambda}}|$ is $O(\lambda^{-\frac{1}{4}})$.

One of the main ingredients is a splitting technique which consists in introducing an artificial “atom” into the state space (if it doesn’t already have one, as in our case) by embedding the original process as a component of a process living in an enlarged state space. In principle, the benefit for having an extended state space with an atom is that the trajectories for the process S_t can be decomposed into a series of i.i.d. parts (i.e. life cycles) corresponding to time intervals $[R_n, R_{n+1})$ where R_n are the return times to the atom. This would allow the integrable functional D_t to be written as a pair of boundary terms plus a sum of i.i.d. random variables (with a random number of terms). For Markov chains, such a technique for embedding an atom was developed independently in [28] and [2], and is referred to as *Nummelin splitting* or merely *splitting*. When it comes to splitting a Markov process, there are different schemes available. In [17], there is a sequence of split processes constructed which contain marginal processes that are arbitrarily close to the original process. The construction in [21] involves a larger state space $\Sigma \times [0, 1] \times \Sigma$, although an exact copy of the original process is embedded as a marginal.

The idea that splitting constructions could be used as a tool to prove certain limit theorems for Markov processes was suggested in an unpublished paper [32].

We use a truncated version of the split process introduced in [21]. The split process is not Markovian itself, but contains an embedded chain (the split resolvent chain) which is Markovian. The life cycles for the process are not completely independent in this construction, since there are correlations between successive life cycles. The details of the construction are explained at the beginning of Section 4, and the discussion is self-contained. The original process S_t , that lives in $\Sigma = \mathbb{T} \times \mathbb{R}$ is embedded as a component of $\tilde{S}_t = (S_t, Z_t) \in \tilde{\Sigma} = \Sigma \times \{0, 1\}$. The process D_t can be written as four boundary terms plus a martingale

$$D_t = (\text{Sum of boundary terms}) + \tilde{M}_t \quad (1.11)$$

$$\tilde{M}_t = \sum_{n=1}^{\tilde{N}_t} \left(\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_n}) + (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_{n+1}}) \right),$$

where \tilde{N}_t is the number of returns to the atom $\Sigma \times 1$ to have occurred before time t , and $\mathfrak{R}^{(\lambda)} : L^\infty(\Sigma) \rightarrow L^\infty(\Sigma)$ is the reduced resolvent of the backwards generator \mathcal{L}_λ . The boundary terms are

$$\int_0^{R_1} dr \frac{dV}{dx}(X_r) - \int_t^{R_{\tilde{N}_t+1}} dr \frac{dV}{dx}(X_r) + (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_1}) - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_{\tilde{N}_t+1}}).$$

The interjection of the telescoping terms $(\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_n})$ removes the correlations between successive life cycles. The fact that the increments of \tilde{M}_t have mean zero with respect to the information known up to time R_n is a consequence of the splitting construction and the fact that the observable $\frac{dV}{dx}$ has mean zero in the equilibrium state $\Psi_{\infty, \lambda}$. The process \tilde{M}_t is a martingale with respect to its own filtration, and this opens the possibility of using Doob's inequality to the bound \tilde{M}_t .

The martingale \tilde{M}_t is a variant of the martingale \tilde{M}' below that is usually employed when studying limit theorems for integral functionals of Markov processes:

$$\tilde{M}'_t = (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_t) - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_0) + D_t, \quad (1.12)$$

The martingale \tilde{M}' has predictable quadratic variation

$$\langle \tilde{M}' \rangle_t = \int_0^t dr \int_{\mathbb{R}} dp' \left((\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(X_r, p') - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(X_r, P_r) \right)^2 \mathcal{J}_\lambda(P_r, p').$$

Using the martingale (1.12) would require showing some decay which is uniform in $\lambda < 1$ for increments $|(\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(x, p') - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(x, p)|$ when $|p|, |p'|$ are large and $|p - p'| = O(1)$. However, it is not clear for us how to obtain the necessary bounds on the resolvent, and our methods are designed for exploiting time-averaging of the process $\frac{dV}{dx}(X_r)$ as suggested by the heuristics in Section 1.2.2. Our technique is based on having bounds for a generalized resolvent $U_h^{(\lambda)} : L^\infty(\Sigma) \rightarrow L^\infty(\Sigma)$ of the form

$$(U_h^{(\lambda)} g)(s) = \mathbb{E}_s^{(\lambda)} \left[\int_0^\infty dt e^{-\int_0^t dr h(S_r)} g(S_t) \right],$$

where h is a non-negative function with compact support, and the function $g = g_\lambda$ essentially has the form

$$g_\lambda(s) = \left| \mathbb{E}_s^{(\lambda)} \left[\int_0^\infty dt t e^{-t} \frac{dV}{dx}(S_t) \right] \right|.$$

The operator $U_h^{(\lambda)}$ arises in the study of recurrence for the Markov process and has been referred to as the *state-modulated resolvent* [23]. Analysis of $U_h^{(\lambda)}$ for our dynamics is contained in [8].

1.2.4 The unit conventions and organization of the article

Throughout the remainder of the article, we will remove units by setting $\beta = a = m = 1$, and picking η such that $\gamma = \frac{1}{2}$. We assume List 1.1 in all theorems, lemmas, etc. unless otherwise stated. Section 2 develops preliminary bounds on the typical energy of the particle. Section 3 contains some analysis of the collision component J_t of the momentum. Section 4 is directed toward proving that $\sup_{0 \leq t \leq T} |D_{\frac{t}{\lambda}}|$ is typically $O(\lambda^{-\frac{1}{4}})$. The proof of the convergence of $\lambda^{\frac{1}{2}} P_{\frac{t}{\lambda}}$ to the Ornstein-Uhlenbeck process \mathbf{p}_t for small λ is found in Section 5, and Section 6 contains miscellaneous proofs that we regarded as mechanical or routine.

2 Typical energy behavior

Functions of the energy $H(x, p) = \frac{1}{2}p^2 + V(x)$ have the advantage of being invariant under the Hamiltonian evolution. This makes energy related quantities a desirable starting point for gaining some control over the typical behavior of the dynamics.

Define the functions $\mathcal{A}_\lambda, \mathcal{V}_{\lambda,n} : \mathbb{T} \times \mathbb{R} \rightarrow \mathbb{R}$ and $\mathcal{E}_\lambda(p) : \mathbb{R} \rightarrow [0, \infty)$ as

$$\begin{aligned} \mathcal{E}_\lambda(p) &= \int_{\mathbb{R}} dp' \mathcal{J}_\lambda(p, p'), \\ \mathcal{A}_\lambda(x, p) &= \int_{\mathbb{R}} dp' \left(2^{\frac{1}{2}} H(x, p')^{\frac{1}{2}} - 2^{\frac{1}{2}} H(x, p)^{\frac{1}{2}} \right) \mathcal{J}_\lambda(p, p'), \\ \mathcal{V}_{\lambda,n}(x, p) &= \int_{\mathbb{R}} dp' \left(2^{\frac{1}{2}} H(x, p')^{\frac{1}{2}} - 2^{\frac{1}{2}} H(x, p)^{\frac{1}{2}} - \frac{\mathcal{A}_\lambda(x, p)}{\mathcal{E}_\lambda(p)} \right)^{2n} \mathcal{J}_\lambda(p, p'). \end{aligned}$$

Here, $\mathcal{E}_\lambda(p)$ is the escape rate from momentum p . We define $\mathcal{A}_\lambda^\pm(s) = \max(\pm \mathcal{A}_\lambda(s), 0)$ to be the positive and negative parts of \mathcal{A}_λ . We will often denote $\mathcal{V}_{1,\lambda}$ as \mathcal{V}_λ . Finally, we also define

$$\begin{aligned} \mathcal{K}_{\lambda,n}(x, p) &= \int_{\mathbb{R}} dp' |H(x, p')^{\frac{1}{2}} - H(x, p)^{\frac{1}{2}}|^n \mathcal{J}_\lambda(p, p'), \\ \mathcal{K}_{\lambda,n}^*(x, p) &= \int_{\mathbb{R}} dp' |H(x, p')^{\frac{1}{2}} - H(x, p)^{\frac{1}{2}}|^n \chi(|p'| > |p|) \mathcal{J}_\lambda(p, p'). \end{aligned}$$

The proof of the proposition below is mostly mechanical and can be found in Section 6.

Proposition 2.1. *There are constants $c, C_n > 0$ such that for λ small enough,*

1. *For all $(x, p) \in \Sigma$, $\mathcal{K}_{\lambda,n}(x, p) \leq C_n(1 + \lambda|p|)^{n+1}$.*
2. *For all $(x, p) \in \Sigma$, $\mathcal{K}_{\lambda,n}^*(x, p) \leq C_n$.*
3. *For all $(x, p) \in \Sigma$, $\mathcal{V}_{\lambda,n}(x, p) \leq C_n(1 + \lambda|p|)$.*

4. For all $(x, p) \in \Sigma$, $\mathcal{A}_\lambda^+(x, p) \leq \frac{C}{1+p^2}$.

5. For all $(x, p) \in \Sigma$, $c \leq \mathcal{V}_\lambda(x, p)$.

Lemma 2.2 states that the energy H_t typically does not go above the scale λ^{-1} over the time interval $[0, \frac{T}{\lambda}]$.

Lemma 2.2. Assume that $\int_\Sigma d\mu(x, p)|p|^n < \infty$. For any $n \in \mathbb{N}$, there exists a $C > 0$ such that

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} (H_r)^{\frac{n}{2}} \right] \leq C \left(\frac{T}{\lambda} \right)^{\frac{n}{2}}$$

for all $T > 0$ and $\lambda < 1$.

Proof. We will work with the process $\mathbf{Q}_t = 2^{\frac{1}{2}} H_t^{\frac{1}{2}}$. The reader should think of \mathbf{Q}_t as being roughly the absolute value of the momentum $|P_t|$. If P_t were a symmetric random walk making steps every unit of time, then the result would hold with $(H_t)^{\frac{1}{2}}$ replaced by $|P_t|$ by Doob's inequality (supposing that the tail distribution of the jumps decays sufficiently fast). The situation for our jump rates should, in principle, be even more accommodating, since the jump rates (1.2) tend to drag a momentum with large absolute value down to a momentum with smaller absolute value. However, for the purposes of this lemma, it is useful to discard the term associated with these large downward jumps in the decomposition (2.2) of \mathbf{Q}_t , since it is less analytically wieldy, and it is not helpful on the time scales $\frac{T}{\lambda}$ for T fixed and $\lambda \ll 1$.

For technical reasons, we partition the time interval $[0, \frac{T}{\lambda}]$ through a sequence of incursion times (ς'_n) into a region of “low” energy. Let $\varsigma_0 = \varsigma'_1 = 0$, and define the stopping times $\varsigma_n, \varsigma'_n$ such that

$$\varsigma'_n = \min\{r \in (\varsigma_{n-1}, \infty) \mid \mathbf{Q}_r \leq \lambda^{-\frac{1}{2}}\}, \quad \varsigma_n = \min\{r \in (\varsigma'_n, \infty) \mid \mathbf{Q}_r \geq 2\lambda^{-\frac{1}{2}}\}.$$

The above definition assumes that $\mathbf{Q}_0 \leq \lambda^{-\frac{1}{2}}$ which is reasonable for $\lambda \ll 1$ by the locality assumption on the initial distribution (2) of List 1.1, but we should take $\varsigma_1 = \varsigma'_1 = 0$ when $\mathbf{Q}_0 > \lambda^{-\frac{1}{2}}$.

Trivially, we have the inequality

$$\sup_{0 \leq r \leq t} \mathbf{Q}_r \leq 2\lambda^{-\frac{1}{2}} + \sup_{0 \leq r \leq t} (\mathbf{Q}_r - \mathbf{Q}_{r-})^+ + \sup_{\varsigma_n \leq t} \sup_{r \in [\varsigma_n, \varsigma'_{n+1} \wedge t]} (\mathbf{Q}_r - \mathbf{Q}_{\varsigma_n})^+. \quad (2.1)$$

The middle term on the right side bounds the over-jumps past the line $2\lambda^{-\frac{1}{2}}$ at the start of the excursions into high energy.

Let $t_n, n > 0$ be the collision times with $t_0 = 0$ and recall that \mathcal{N}_t is the number of collisions up to time t . We can write \mathbf{Q}_t as

$$\mathbf{Q}_t = \mathbf{Q}_0 + \mathbf{m}_t + \mathbf{m}'_t + \int_0^t dr \mathcal{A}^+(S_r) - \sum_{j=1}^{\mathcal{N}_t} \frac{\mathcal{A}_\lambda^-(S_{t_j^-})}{\mathcal{E}_\lambda(S_{t_j^-})}, \quad (2.2)$$

where the process \mathbf{m}_t is defined as

$$\mathbf{m}_t := \sum_{j=1}^{\mathcal{N}_t} \Delta_j \quad \text{for} \quad \Delta_j = \mathbf{Q}_{t_j} - \mathbf{Q}_{t_j^-} - \frac{\mathcal{A}_\lambda(S_{t_j^-})}{\mathcal{E}_\lambda(S_{t_j^-})},$$

and where \mathbf{m}'_t is the difference

$$\mathbf{m}'_t = \sum_{j=1}^{\mathcal{N}_t} \frac{\mathcal{A}_\lambda^+(S_{t_j^-})}{\mathcal{E}_\lambda(S_{t_j^-})} - \int_0^t dr \mathcal{A}_\lambda^+(S_r).$$

The processes \mathbf{m}_t and \mathbf{m}'_t are martingales with respect to the full filtration \mathcal{F}_t . For \mathbf{m}_t , the increments Δ_n have mean zero with respect to the information $\mathcal{F}_{\tau_j^-}$ and that $\mathcal{N}_r = \mathcal{N}_{r-} + 1$.

The process \mathbf{m}_t is a martingale, since the terms $\frac{\mathcal{A}_\lambda^+(S_r)}{\mathcal{E}_\lambda(S_r)}$ in the sum occur with Poisson rate $\mathcal{E}_\lambda(r)$. Moreover, the predictable quadratic variations corresponding to the martingales have the forms

$$\langle \mathbf{m} \rangle_t = \int_0^t dr \mathcal{V}_\lambda(S_r) \quad \text{and} \quad \langle \mathbf{m}' \rangle_t = \int_0^t dr \frac{(\mathcal{A}_\lambda^+(S_r))^2}{\mathcal{E}_\lambda(S_r)}.$$

The third term on the right side of (2.1) is smaller than

$$\begin{aligned} \sup_{\varsigma_n \leq \frac{T}{\lambda}} \sup_{t \in [\varsigma_n, \varsigma'_{n+1} \wedge \frac{T}{\lambda})} (\mathbf{Q}_t - \mathbf{Q}_{\varsigma_n})^+ &\leq \sup_{\varsigma_n \leq \frac{T}{\lambda}} \sup_{t \in [\varsigma_n, \varsigma'_{n+1} \wedge \frac{T}{\lambda})} \left(\mathbf{m}_t + \mathbf{m}'_t - \mathbf{m}_{\varsigma_n} - \mathbf{m}'_{\varsigma_n} + \int_{\varsigma_n}^t dr \mathcal{A}^+(S_r) \right)^+ \\ &\leq 2 \sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}_t| + 2 \sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}'_t| + \int_0^{\frac{T}{\lambda}} \mathcal{A}^+(S_r) \chi(\mathbf{Q}_r \geq \lambda^{-\frac{1}{2}}) \\ &\leq 2 \sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}_t| + 2 \sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}'_t| + 2CT. \end{aligned} \quad (2.3)$$

For the first inequality, we have thrown away the term $-\sum_{j=1}^{\mathcal{N}_t} \frac{\mathcal{A}_\lambda^-(S_{t_j^-})}{\mathcal{E}_\lambda(S_{t_j^-})}$, since it is strictly negative. The second inequality is the triangle inequality with $\sup_{0 \leq s, r \leq t} |f_r - f_s| \leq 2 \sup_{0 \leq r \leq t} |f_r|$ for $f = \mathbf{m}, \mathbf{m}'$, and uses the fact that $\mathbf{Q}_r \geq \lambda^{-\frac{1}{2}}$ during the excursion intervals $[\varsigma_n, \varsigma'_{n+1} \wedge \frac{T}{\lambda})$. The third inequality is a consequence of Part (4) of Proposition 2.1 which gives a C such that

$$\mathcal{A}^+(X_r, P_r) \chi(\mathbf{Q}_r \geq \lambda^{-\frac{1}{2}}) \leq C \frac{1}{1 + P_r^2} \chi(\mathbf{Q}_r \geq \lambda^{-\frac{1}{2}}) \leq \frac{C}{1 + \frac{1}{2} \lambda^{-1}} \leq 2C\lambda.$$

The second inequality is for λ small enough so that $4 \sup_x V(x) \leq \lambda^{-1}$.

Combining (2.2) with (2.3) and using the triangle inequality, then

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{n}} &\leq 2CT + \mathbb{E}^{(\lambda)} [\mathbf{Q}_0^n]^{\frac{1}{n}} + \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} ((\mathbf{Q}_t - \mathbf{Q}_{t-})^+)^n \right]^{\frac{1}{n}} \\ &\quad + 2\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}_t|^n \right]^{\frac{1}{n}} + 2\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}'_t|^n \right]^{\frac{1}{n}}. \end{aligned} \quad (2.4)$$

We will now give bounds for each of the terms on the right side. The goal is to show that

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{n}} = O \left(\lambda^{-\frac{1}{2}} + \sum_m \left(\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{n}} \right)^{\alpha_m} \right), \quad (2.5)$$

where $\alpha_m \geq 2$ and the sum over m is finite. The above would imply that $\mathbb{E}^{(\lambda)} [\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n]$ is $O(\lambda^{-\frac{n}{2}})$, which is the statement of the lemma.

For the second term on the right side of (2.4),

$$\mathbb{E}^{(\lambda)}[\mathbf{Q}_0^n] = \int_S d\mu(x, p) (p^2 + 2V(x))^{\frac{n}{2}} < \infty,$$

and the right side is finite by our assumption on the initial measure in List 1.1. For the third term on the right side of (2.4). Using that $(\sup_m a_m)^2 \leq a_m^2$ and Jensen's inequality, we have the first inequality below

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} |\mathbf{Q}_r - \mathbf{Q}_{r-}|^n \right]^{\frac{1}{n}} &\leq \mathbb{E}^{(\lambda)} \left[\sum_{m=1}^{\mathcal{N}_{\frac{T}{\lambda}}} ((\mathbf{Q}_{t_m} - \mathbf{Q}_{t_m-})^+)^{2n} \right]^{\frac{1}{2n}} = \mathbb{E}^{(\lambda)} \left[\sum_{m=1}^{\mathcal{N}_{\frac{T}{\lambda}}} \frac{\mathcal{K}_{\lambda, 2n}^*(X_{t_m-}, P_{t_m-})}{\mathcal{E}_{\lambda}(P_{t_m-})} \right]^{\frac{1}{2n}} \\ &= \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dt \mathcal{K}_{\lambda, 2n}^*(X_t, P_t) \right]^{\frac{1}{2n}} \leq C^{\frac{1}{2n}} \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dt (1 + \lambda \mathbf{Q}_t) \right]^{\frac{1}{2n}} \\ &\leq C^{\frac{1}{2n}} T^{\frac{1}{2}} \lambda^{-\frac{1}{2n}} + C^{\frac{1}{2n}} T^{\frac{1}{2n}} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{2n^2}}. \end{aligned}$$

The first equality uses that

$$\mathbb{E} \left[((\mathbf{Q}_r - \mathbf{Q}_{r-})^+)^{2n} \mid \mathcal{F}_{r-}, \mathcal{N}_r = \mathcal{N}_{r-} + 1 \right] = \frac{\mathcal{K}_{\lambda, 2n}^+(X_r, P_r)}{\mathcal{E}_{\lambda}(P_r)},$$

and the second equality uses that the times t_n occur with Poisson rate $\mathcal{E}_{\lambda}(P_r)$. The second inequality is from Part (2) of Proposition 2.1. The right-most term above has the form of (2.5).

For the fourth term on the right side of (2.4), we can apply Doob's maximal inequality to get the first inequality below

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}_t|^n \right]^{\frac{1}{n}} &\leq \frac{n}{n-1} \mathbb{E}^{(\lambda)} [|\mathbf{m}_{\frac{T}{\lambda}}|^n]^{\frac{1}{n}} \leq C' \mathbb{E}^{(\lambda)} [(\langle \mathbf{m} \rangle_{\frac{T}{\lambda}})^{\frac{n}{2}}]^{\frac{1}{n}} + C' \mathbb{E}^{(\lambda)} \left[\sum_{m=1}^{\mathcal{N}_{\frac{T}{\lambda}}} |\Delta_m|^n \right]^{\frac{1}{n}} \\ &= C' \mathbb{E}^{(\lambda)} \left[\left(\int_0^{\frac{T}{\lambda}} dr \mathcal{V}_{\lambda}(S_r) \right)^{\frac{n}{2}} \right]^{\frac{1}{n}} + C' \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr \mathcal{V}_{\lambda, n}(S_r) \right]^{\frac{1}{n}} \\ &\leq C'' \mathbb{E}^{(\lambda)} \left[\left(\int_0^{\frac{T}{\lambda}} dr (1 + \lambda \mathbf{Q}_r) \right)^{\frac{n}{2}} \right]^{\frac{1}{n}} + C'' \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr (1 + \lambda \mathbf{Q}_r) \right]^{\frac{1}{n}} \\ &\leq C'' (T^{\frac{1}{2}} \lambda^{-\frac{1}{2}} + T^{\frac{1}{n}} \lambda^{-\frac{1}{n}}) + C'' T^{\frac{1}{2}} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{2n}} + C'' T^{\frac{1}{n}} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \mathbf{Q}_t^n \right]^{\frac{1}{n^2}}. \end{aligned}$$

The second inequality is for some $C' > 0$ by Rosenthal's inequality (see e.g. [11], Lemma 2.1). The third inequality is Part (3) of Proposition 2.1. We have combined the constants at each step. Thus we have made dressed our bound in the form (2.5).

The last term in 2.4 is bounded similarly to $\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\mathbf{m}_t|^n \right]^{\frac{1}{n}}$.

□

The following lemma gives a bound for the expected amount of time over an interval $[0, \frac{T}{\lambda}]$ that the process S_t spends with energy smaller than $\lambda^{-\varrho}$ for $\varrho \in [0, 1)$. The result is similar to Lemma 5.1 of [6], and we modify the proof for our purposes here.

Lemma 2.3. Define $\mathbf{T}_t = \lambda \int_0^t dr \chi(H_r \leq \epsilon \lambda^{-\varrho})$ for $0 \leq \varrho \leq 1$. For any fixed $T > 0$, there is a $C > 0$ such that for small enough λ and all $\epsilon \geq \lambda^\varrho$,

$$\mathbb{E}^{(\lambda)}[\mathbf{T}_{\frac{T}{\lambda}}] \leq C \epsilon^{\frac{1}{2}} \lambda^{\frac{1-\varrho}{2}}.$$

Proof.

For $b > 0$, let γ be the minimum of the hitting time that H_t jumps above $b\lambda^{-2}$ and the final time $\frac{T}{\lambda}$. We have the following inequalities,

$$\begin{aligned} \mathbb{E}^{(\lambda)}[\mathbf{T}_{\frac{T}{\lambda}}] &\leq T \mathbb{P}^{(\lambda)}\left[\sup_{0 \leq r \leq \frac{T}{\lambda}} H_r > b\lambda^{-2}\right] + \mathbb{E}^{(\lambda)}[\mathbf{T}_\gamma] \\ &\leq T \frac{\lambda^4}{b^2} T \mathbb{E}^{(\lambda)}\left[\left(\sup_{0 \leq r \leq \frac{T}{\lambda}} H_r\right)^2\right] + \mathbb{E}^{(\lambda)}[\mathbf{T}_\gamma] \leq CT^3 \frac{\lambda^2}{b^2} + \mathbb{E}^{(\lambda)}[\mathbf{T}_\gamma], \end{aligned}$$

where the second inequality is Chebyshev's and the $C > 0$ in the third is from Lemma 2.2. With the restriction $\epsilon \geq \lambda^\varrho$, the term following the last inequality decays faster than $\epsilon^{\frac{1}{2}} \lambda^{\frac{1-\varrho}{2}}$ as $\lambda \rightarrow 0$, so we can focus our study to $\mathbb{E}^{(\lambda)}[\mathbf{T}_\gamma]$. The energy process $H_t = H(X_t, P_t)$ behaves as a submartingale for time periods in which $H_t \leq b\lambda^{-2}$ for small enough $b > 0$. More precisely, there exists $0 \leq b' \leq 1$, $\underline{\sigma} > 0$ such that for all $\lambda < 1$ and all (x, p) with $H(x, p) \leq b'\lambda^{-2}$,

$$\begin{aligned} \mathbf{q}_1^{(\lambda)}(x, p) &:= \frac{d}{dt} \mathbb{E}_{(x, p)}^{(\lambda)}[H_t] \Big|_{t=0} = \frac{1}{2} \int_{\mathbb{R}} dp' ((p')^2 + V(x) - p^2 + V(x)) \mathcal{J}_\lambda(p, p') \\ &= \frac{1}{2} \int_{\mathbb{R}} dp' ((p')^2 - p^2) \mathcal{J}_\lambda(p, p') \geq \underline{\sigma}. \end{aligned} \quad (2.6)$$

From (2.6), we have for all $m \geq 1$, $H(x, p) \leq b'\lambda^{-2}$, $\lambda < 1$ that

$$\begin{aligned} \mathbf{q}_m^{(\lambda)}(x, p) &= \int_{\mathbb{R}} dp' \left(\left(\frac{1}{2}(p')^2 + V(x) \right)^m - \left(\frac{1}{2}p^2 + V(x) \right)^m \right) \mathcal{J}_\lambda(p, p') \\ &\geq m \underline{\sigma} H^{m-1}(x, p), \end{aligned} \quad (2.7)$$

where $\mathbf{q}_m^{(\lambda)}(x, p) := \frac{d}{dt} \mathbb{E}_{(x, p)}^{(\lambda)}[H_t^m] \Big|_{t=0}$. The inequality (2.7) follows from (2.6), since $f(y) = |p|^m$ is convex, and thus

$$f(Y_1) - f(Y_0) \geq (Y_1 - Y_0) f'(Y_0)$$

for $Y_1 = \frac{1}{2}(p')^2 + V(x)$ and $Y_0 = \frac{1}{2}p^2 + V(x)$.

Notice that the value $\mathbf{q}_m^{(\lambda)}(S_t)$ is the derivative of the predictable part of the semimartingale decomposition of the process H_t^m . In other terms,

$$H_t^m - \int_0^t dr \mathbf{q}_m^{(\lambda)}(S_r)$$

is a martingale. In addition to the lower bounds in (2.7), there are upper bounds

$$\mathbf{q}_m^{(\lambda)}(x, p) \leq \bar{\sigma}_m (1 + H^{m-1}(x, p)) \quad (2.8)$$

which hold for some constants $\bar{\sigma}_m$ and all $\lambda < 1$ and (x, p) with $H(x, p) \leq \lambda^{-2}$.

Using the above observations, there is a useful submartingale which is “close” to $H_t^{\frac{1}{2}}$. Let the $b > 0$ defining the stopping time γ be the b' chosen to ensure the condition (2.6). There exists a $c > 0$ such that for all λ small enough,

$$\kappa_t = H_t^{\frac{1}{2}} + c\lambda H_t^{\frac{3}{2}}$$

is a submartingale over the time interval $t \in [0, \gamma]$. To see that κ_t is a submartingale up to time γ , first notice that the predictable component $\int_0^t dr \mathbf{q}_{\frac{3}{2}}^{(\lambda)}(S_r)$ in the semimartingale decomposition of $H_t^{\frac{3}{2}}$ increases with rate $\underline{\sigma} H_t^{\frac{1}{2}}$ by (2.7). Moreover, the predictable component of the semimartingale decomposition of $H_t^{\frac{1}{2}}$ is $2^{-\frac{1}{2}} \int_0^t dr \mathcal{A}_\lambda(X_t, P_t)$, and the negative part of the function \mathcal{A}_λ satisfies the inequality $\mathcal{A}_\lambda^-(x, p) \leq C\lambda H^{\frac{1}{2}}(x, p)$ for some $C > 0$ by Part (2) of Proposition 4.15 and the elementary inequality $|p| \leq 2^{\frac{1}{2}} H^{\frac{1}{2}}(x, p)$. Thus we can choose $c = \frac{C}{\underline{\sigma}}$ to ensure that κ_t is a submartingale in the specified time interval.

Set $\varsigma_0 = \varsigma'_1 = 0$, and define the stopping times $\varsigma_n, \varsigma'_n \leq \gamma$ such that

$$\varsigma'_n = \min\{r \in (\varsigma_{n-1}, \infty) \mid H_r \leq \epsilon\lambda^{-e}\}, \quad \varsigma_n = \min\{r \in (\varsigma'_n, \infty) \mid H_r \geq 4\epsilon\lambda^{-e}\}.$$

Let \mathbf{n}_γ be the number ς'_n 's less than γ . In other words, \mathbf{n}_γ is one more than the number of up-crossings of H_r from λ^{-e} to $4\lambda^{-e}$ which have been completed by time γ . Let \mathbf{n}'_γ be defined similarly as one plus the number of crossings of κ_t from $\frac{3}{2}\epsilon^{\frac{1}{2}}\lambda^{-\frac{e}{2}}$ to $2\epsilon^{\frac{1}{2}}\lambda^{-\frac{e}{2}}$. For $\lambda < (\frac{1}{2c})^{\frac{1}{1-e}}$, then we have both of the implications

$$H_t \geq 4\epsilon\lambda^{-e} \implies \kappa_t \geq 2\epsilon^{\frac{1}{2}}\lambda^{-\frac{e}{2}} \quad \text{and} \quad H_t \leq \epsilon\lambda^{-e} \implies \kappa_t \leq \frac{3}{2}\epsilon^{\frac{1}{2}}\lambda^{-\frac{e}{2}},$$

and hence $\mathbf{n}'_\gamma \geq \mathbf{n}_\gamma$. The definitions give us the almost always inequality

$$\mathbf{T}_t \leq \lambda \sum_{n=1}^{\mathbf{n}_t} \varsigma_n - \varsigma'_n.$$

Next, observe that

$$\lambda^{-1} \mathbb{E}^{(\lambda)}[\mathbf{T}_\gamma] \leq \mathbb{E}^{(\lambda)}\left[\sum_{n=1}^{\mathbf{n}_\gamma} \varsigma_n - \varsigma'_n\right] \leq \mathbb{E}^{(\lambda)}[\mathbf{n}_\gamma] \sup_{n \in \mathbb{N}} \mathbb{E}^{(\lambda)}[\varsigma_n - \varsigma'_n \mid n \leq \mathbf{n}_\gamma]. \quad (2.9)$$

We now have our upper bound in terms of the expectation for the number of up-crossings \mathbf{n}_t and the expectation for the time of a single up-crossing $\varsigma_n - \varsigma'_n$ conditioned on the event $n \leq \mathbf{n}_t$. By the observation above $\mathbb{E}^{(\lambda)}[\mathbf{n}_\gamma] \leq \mathbb{E}^{(\lambda)}[\mathbf{n}'_\gamma]$. By the submartingale up-crossing inequality [5],

$$\begin{aligned} \mathbb{E}^{(\lambda)}[\mathbf{n}'_\gamma] &\leq \frac{\mathbb{E}^{(\lambda)}[\kappa_\gamma]}{2\epsilon^{\frac{1}{2}}\lambda^{-\frac{e}{2}} - \epsilon^{\frac{1}{2}}\frac{3}{2}\lambda^{-\frac{e}{2}}} = 2\epsilon^{-\frac{1}{2}}\lambda^{\frac{e}{2}} \mathbb{E}^{(\lambda)}\left[H_\gamma^{\frac{1}{2}} + c\lambda H_\gamma^{\frac{3}{2}}\right] \leq 2\epsilon^{-\frac{1}{2}}\lambda^{\frac{e}{2}} \left(\mathbb{E}^{(\lambda)}[H_\gamma]^{\frac{1}{2}} + c\lambda \mathbb{E}[H_\gamma^{\frac{3}{2}}]\right) \\ &\leq 2\epsilon^{-\frac{1}{2}}\lambda^{\frac{e}{2}} \left(\mathbb{E}^{(\lambda)}[H_0] + \lambda^{-1}T\bar{\sigma}_1\right)^{\frac{1}{2}} + 2c\epsilon^{-\frac{1}{2}}\lambda^{1+\frac{e}{2}} \left(\mathbb{E}^{(\lambda)}[H_0] + \lambda^{-\frac{3}{2}+\frac{e}{2}}T^{\frac{3}{2}}\bar{\sigma}_{\frac{3}{2}}\right) \\ &< 4\epsilon^{-\frac{1}{2}}\bar{\sigma}_1^{\frac{1}{2}}T^{\frac{1}{2}}\lambda^{\frac{e-1}{2}}, \end{aligned} \quad (2.10)$$

where the last inequality is for λ small enough. The second inequality is Jensen's, and the third uses that $\gamma \leq \frac{T}{\lambda}$ and the bound (2.8) for the derivatives of the predictable components of the semimartingales H_t and $H_t^{\frac{3}{2}}$.

We now focus on the expectation of the incursions $\varsigma_n - \varsigma'_n$. Whether or not the event $n \leq \mathbf{n}_t$ occurred will be known at time ς'_n , so

$$\sup_{n \in \mathbb{N}} \mathbb{E}^{(\lambda)}[\varsigma_n - \varsigma'_n \mid n \leq \mathbf{n}_t] \leq \sup_{n \in \mathbb{N}, \omega \in \mathcal{F}_{\varsigma'_n}} \mathbb{E}^{(\lambda)}[\varsigma_n - \varsigma'_n \mid \mathcal{F}_{\varsigma'_n}] = \sup_{H(s) \leq \lambda^{-\varrho}} \mathbb{E}_s^{(\lambda)}[\varsigma_1].$$

By (2.6), then

$$\sigma \mathbb{E}_s^{(\lambda)}[\varsigma_1] \leq \mathbb{E}_s^{(\lambda)}\left[\int_0^{\varsigma_1} dr \mathbf{q}_1^{(\lambda)}(X_r, P_r)\right] = \mathbb{E}_s^{(\lambda)}[H_{\varsigma_1} - H_0], \quad (2.11)$$

where the equality holds by the optional sampling theorem, since $\int_0^t dr \mathbf{q}_1^{(\lambda)}(X_r, P_r)$ is the predictable part of the semimartingale decomposition for $H_t - H_0$. This application of the optional sampling theorem is legal, since ς_1 is almost surely finite, and

$$\mathbb{E}_s^{(\lambda)}[H_r \chi(r < \varsigma_1)] \leq \lambda^{-\frac{1}{2}} \mathbb{P}_s^{(\lambda)}[r < \tau_1] \longrightarrow 0 \quad \text{as } r \longrightarrow \infty.$$

Continuing with the right side of (2.11),

$$\begin{aligned} \mathbb{E}_s^{(\lambda)}[H_{\varsigma_1} - H_0] &= \mathbb{E}_s^{(\lambda)}[H_{\varsigma_1^-} - H_0] + \frac{1}{2} \mathbb{E}_s^{(\lambda)}[P_{\varsigma_1}^2 - P_{\varsigma_1^-}^2] \\ &\leq 2\epsilon \lambda^{-\varrho} + 2\mathbb{E}_s^{(\lambda)}[\Delta^2] = 2\epsilon \lambda^{-\varrho} + O(\lambda^0) \leq c\epsilon \lambda^{-\varrho}. \end{aligned}$$

where $\Delta = P_{\varsigma_1} - P_{\varsigma_1^-}$. For the first inequality, we have used that $P_{\varsigma_1^-}^2 \leq H_{\varsigma_1^-} \leq \epsilon \lambda^{-\varrho}$, and the inequality $(x + y)^2 \leq 2(x^2 + y^2)$. The last inequality holds for some $c > 0$ by our restriction $\epsilon \geq \lambda^{\varrho}$. The term $\mathbb{E}_s^{(\lambda)}[\Delta^2]$ is uniformly bounded for $\lambda \leq 1$, since by nested conditional expectations $\mathbb{E}_s^{(\lambda)}[\Delta^2] = \mathbb{E}_s^{(\lambda)}[\mathbb{E}^{(\lambda)}[\Delta^2 \mid \mathcal{F}_{\varsigma_1^-}]]$ and

$$\begin{aligned} \mathbb{E}^{(\lambda)}[\Delta^2 \mid \mathcal{F}_{\varsigma_1^-}] &= \frac{\int_{H(X_{\varsigma_1^-}, p') \geq \epsilon \lambda^{-\varrho}} dp' (p' - P_{\varsigma_1^-})^2 \mathcal{J}_{\lambda}(P_{\varsigma_1^-}, p')}{\int_{H(X_{\varsigma_1^-}, p') \geq \epsilon \lambda^{-\varrho}} dp' \mathcal{J}_{\lambda}(P_{\varsigma_1^-}, p')} \\ &\leq \sup_{\lambda < 1} \sup_{H(x, p) \leq \epsilon \lambda^{-\varrho}} \frac{\int_{H(x, p') \geq \epsilon \lambda^{-\varrho}} dp' (p' - p)^2 \mathcal{J}_{\lambda}(p, p')}{\int_{H(x, p') \geq \epsilon \lambda^{-\varrho}} dp' \mathcal{J}_{\lambda}(p, p')} < \infty. \end{aligned}$$

The equality relies on the strong Markov property, since the distribution for Δ is independent of $\mathcal{F}_{\varsigma_1^-}$ when given $S_{\varsigma_1^-}$. The last expression is finite by Part (7) of Proposition 3.1.

Putting our results for $\mathbb{E}^{(\lambda)}[\mathbf{n}_{\gamma}]$ and $\sup_{n \in \mathbb{N}} \mathbb{E}^{(\lambda)}[\varsigma_n - \varsigma'_n \mid n \leq \mathbf{n}_{\gamma}]$ together,

$$\mathbb{E}^{(\lambda)}[\mathbf{T}_{\gamma}] \leq \lambda \mathbb{E}^{(\lambda)}[\mathbf{n}_{\gamma}] \sup_{n \in \mathbb{N}} \mathbb{E}^{(\lambda)}[\varsigma_n - \varsigma'_n \mid n \leq \mathbf{n}_{\gamma}] \leq 4c\epsilon^{\frac{1}{2}} \bar{\sigma}^{\frac{1}{2}} T^{\frac{1}{2}} \lambda^{\frac{1-\varrho}{2}}.$$

This completes the proof. □

The process $\mathbf{A}_t = \int_0^t dr \mathcal{A}_{\lambda}(S_r)$ is the predictable part of the semimartingale decomposition of $\mathbf{Q}_t = 2^{\frac{1}{2}} H_t^{\frac{1}{2}}$. Moreover, $\mathbf{A}_t^{\pm} = \int_0^t dr \mathcal{A}_{\lambda}^{\pm}(S_r)$ are the increasing/decreasing parts in the Jordan decomposition of \mathbf{A}_t . The martingale part \mathbf{M}_t of the semimartingale decomposition of \mathbf{Q}_t has predictable quadratic variation

$$\langle \mathbf{M} \rangle_t = \int_0^t dr \mathcal{V}_{\lambda}(S_r).$$

Lemma 2.4. *There exists $C, c > 0$ such that for all $t \in [0, \frac{T}{\lambda}]$, $\lambda < 1$ and s in a given compact subset of Σ ,*

$$\mathbb{E}_s^{(\lambda)}[\mathbf{A}_t^+] \geq -C + ct^{\frac{1}{2}}.$$

Proof. Since \mathbf{Q}_t and \mathbf{A}_t^- are positive and $\mathbf{A}_t^+ \geq 0$ is increasing, we have the following deterministic inequalities

$$\mathbf{A}_t^+ = \mathbf{Q}_t - \mathbf{Q}_0 - \mathbf{M}_t + \mathbf{A}_t^- \geq -\mathbf{Q}_0 + \sup_{0 \leq r \leq t} -\mathbf{M}_r \geq -\mathbf{Q}_0 + \mathbf{M}_t^-.$$

Taking the expectation of both side gives.

$$\mathbb{E}_s^{(\lambda)}[t^{-\frac{1}{2}}\mathbf{A}_t^+] \geq -2^{\frac{1}{2}}t^{-\frac{1}{2}}H^{\frac{1}{2}}(s) + t^{-\frac{1}{2}}\mathbb{E}_s^{(\lambda)}[\mathbf{M}_t^-].$$

Since \mathbf{M}_t has mean zero, we have the equality below

$$2\mathbb{E}_s^{(\lambda)}[\mathbf{M}_t^-] = \mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|] \geq \frac{\mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^2]^2}{\mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^3]},$$

and the inequality is Cauchy-Schwarz. However,

$$\mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^2] = \mathbb{E}_s^{(\lambda)}\left[\int_0^t dr \mathcal{V}_\lambda(S_r)\right] \geq ct,$$

where $c > 0$ is from Part (5) of Proposition 2.1. For the first inequality below, we use Rosenthal's inequality (see e.g. [11], Lemma 2.1) to produce a $C > 0$ such that

$$\begin{aligned} \mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^3] &\leq C\mathbb{E}_s^{(\lambda)}[\langle \mathbf{M}_t \rangle^{\frac{3}{2}}] + C\mathbb{E}_s^{(\lambda)}\left[\sum_{n=1}^{\mathcal{N}_t} |\mathbf{M}_{t_n} - \mathbf{M}_{t_n^-}|^3\right] \\ &= C\mathbb{E}_s^{(\lambda)}[\langle \mathbf{M}_t \rangle^{\frac{3}{2}}] + C\mathbb{E}_s^{(\lambda)}\left[\sum_{n=1}^{\mathcal{N}_t} |\mathbf{Q}_{t_n} - \mathbf{Q}_{t_n^-}|^3\right] = C\mathbb{E}_s^{(\lambda)}\left[\left(\int_0^t dr \mathcal{V}_\lambda(S_r)\right)^{\frac{3}{2}} + \int_0^t dr \mathcal{K}_{\lambda,3}(S_r)\right] \\ &\leq C'\mathbb{E}_s^{(\lambda)}\left[\left(\int_0^t dr (1 + \lambda \mathbf{Q}_r)\right)^{\frac{3}{2}} + \int_0^t dr (1 + \lambda \mathbf{Q}_r)^4\right] \leq C''t^{\frac{3}{2}}, \end{aligned}$$

where t_n are the collision times and \mathcal{N}_t is the number of collisions up to time t . The first equality uses that \mathbf{Q}_t and \mathbf{M}_t differ by a continuous process and thus have the same jumps. The second inequality is for some $C' > 0$ by Part (1) and (3) of Proposition 2.1 along with the relation $|p| \leq 2^{\frac{1}{2}}H^{\frac{1}{2}}(x, p)$. The last inequality is by Lemma 2.2, and C'' is independent of $\lambda < 1$.

Putting our results together

$$\mathbb{E}_s^{(\lambda)}[\mathbf{A}_t^+] \geq -2^{\frac{1}{2}}H^{\frac{1}{2}}(s) + \frac{\mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^2]^2}{\mathbb{E}_s^{(\lambda)}[|\mathbf{M}_t|^3]} \geq -2^{\frac{1}{2}}H^{\frac{1}{2}}(s) + \frac{c^2}{C''}t^{\frac{1}{2}}.$$

This gives the result. □

3 Limiting behavior for the jump process

In this section, we push the study of the jump processes J_t in (1.3) as far as we can without any analysis of the drift D_t . J_t can be written as a sum of a martingale M_t and a predictable part as

$$J_t = M_t + \int_0^t dr \mathcal{D}_\lambda(P_r) \quad \text{for} \quad \mathcal{D}_\lambda(p) = \int_{\mathbb{R}} dp' (p' - p) \mathcal{J}_\lambda(p, p').$$

In order to write an expression for the predictable quadratic variation $\langle M \rangle_t$ in terms of the jump rates \mathcal{J}_λ , let us define for $m \in \mathbb{R}^+$,

$$\Pi_{\lambda,m}(p) = \int_{\mathbb{R}} dp' (p' - p)^m \mathcal{J}_\lambda(p, p'), \quad \text{and} \quad \mathcal{Q}_\lambda(p) = \Pi_{\lambda,2}(p) - \frac{\mathcal{D}_\lambda^2(p)}{\mathcal{E}_\lambda(p)}.$$

Note that \mathcal{E}_λ , \mathcal{D}_λ are just special cases of $\Pi_{\lambda,m}$ for $m = 0$ and $m = 1$, respectively. The predictable quadratic variation for M_t can then be written in the closed form

$$\langle M \rangle_t = \int_0^t dr \mathcal{Q}_\lambda(P_r).$$

The proof of the following proposition is also mostly mechanical and can be found in Section 6.

Proposition 3.1. *There are constants $C, C_m, c > 0$ such that for all $\lambda \leq 1$ and all $p \in \mathbb{R}$*

1. $\frac{1}{8(\lambda+1)} \leq \mathcal{E}_\lambda(p) \leq \frac{1}{8(\lambda+1)} (1 + C \min(\lambda|p|, \lambda^2 p^2))$ and $\lambda|p| \leq C \mathcal{E}_\lambda(p)$.
2. $|\mathcal{D}_\lambda(p) + \frac{\lambda p}{2(\lambda+1)^2}| \leq C \lambda^2 p^2$.
3. $|\mathcal{D}_\lambda(p) + \frac{2\lambda p}{\lambda+1} \mathcal{E}_\lambda(p)| \leq C$.
4. $|\mathcal{Q}_\lambda(p) - \frac{1}{(\lambda+1)^3}| \leq C \min(\lambda|p|, \lambda^2 p^2)$.
5. $0 < c \leq \mathcal{Q}_\lambda(p) \leq C \mathcal{E}_\lambda(p)$.
6. $\Pi_{\lambda,2m}(p) \leq C_m (1 + \lambda|p|)^{2m+1}$.
- 7.

$$\sup_{\lambda < 1} \sup_{|p| \leq b \leq \lambda^{-1}} \frac{\int_{|p'| \geq b} dp' (|p'| - b)^m \mathcal{J}_\lambda(p, p')}{\int_{|p'| \geq b} dp' \mathcal{J}_\lambda(p, p')} \leq C_m,$$

$$\sup_{\lambda < 1} \sup_{|p| > \frac{1}{\lambda}} \frac{\int_{[-\frac{1}{\lambda}, \frac{1}{\lambda}]} dp' (|p'| - \frac{1}{\lambda})^m \mathcal{J}_\lambda(p, p')}{\int_{[-\frac{1}{\lambda}, \frac{1}{\lambda}]} dp' \mathcal{J}_\lambda(p, p')} \leq C_m.$$

Next, we show that the drift rate $\mathcal{D}_\lambda(p)$ can be effectively replaced by the linear form $-\frac{1}{2}\lambda p$. Let P'_t be the solution to the integral equation

$$P'_t = P_0 + \int_0^t dr \frac{dV}{dx}(X_r) - \frac{1}{2}\lambda \int_0^t dr P'_r + M_t. \quad (3.1)$$

Also let us denote $P_t^{(\lambda)} = \lambda^{\frac{1}{2}} P_t^{(\lambda)} \frac{1}{\lambda}$ and $P_t^{(\lambda),'} = \lambda^{\frac{1}{2}} P'_t \frac{1}{\lambda}$.

Lemma 3.2. *For all fixed $T > 0$, and as $\lambda \rightarrow 0$, the difference between $P_t^{(\lambda),\prime}$ and $P_t^{(\lambda)}$ is small in the sense*

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |P_t^{(\lambda),\prime} - P_t^{(\lambda)}| \right] = O(\lambda^{\frac{1}{2}}).$$

Proof. Since P_t satisfies the integral equation

$$P_t = P_0 + \int_0^t dr \frac{dV}{dx}(X_r) - \int_0^t dr \mathcal{D}_\lambda(P_r) + M_t,$$

and P_t' satisfies (3.1), we have the identity

$$P_t^{(\lambda),\prime} - P_t^{(\lambda)} = \int_0^t dr R_\lambda(P_r^{(\lambda),\prime}) + \frac{1}{2} \int_0^t dr e^{\frac{1}{2}(t-r)} \int_0^r ds R_\lambda(P_s^{(\lambda),\prime}) = \int_0^t dr e^{\frac{1}{2}(t-r)} R_\lambda(P_r^{(\lambda),\prime}),$$

where $R_\lambda(p) = \lambda^{-\frac{1}{2}} D_\lambda(\frac{p}{\lambda^{\frac{1}{2}}}) + \frac{1}{2}p$. Thus, we have the first inequality below

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |P_t^{(\lambda),\prime} - P_t^{(\lambda)}| \right] &\leq 2(e^{\frac{1}{2}T} - 1) \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |R_\lambda(P_t^{(\lambda)})| \right] \\ &\leq 2\lambda^{\frac{1}{2}} C e^{\frac{1}{2}T} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |\lambda^{\frac{1}{2}} P_t^{\frac{1}{\lambda}}| \right] + 2\lambda C e^{\frac{1}{2}T} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |\lambda^{\frac{1}{2}} P_t^{\frac{1}{\lambda}}|^2 \right] = O(\lambda^{\frac{1}{2}}). \end{aligned} \quad (3.2)$$

The second inequality follows by Part (2) of Proposition 3.1. By bounding $|p_r| \leq 2^{\frac{1}{2}} H_r^{\frac{1}{2}}$ and applying Lemma 2.2, the expectations on the right side are uniformly bounded for $\lambda < 1$. Thus, the left side is $O(\lambda^{\frac{1}{2}})$. \square

The following lemma gives a central limit theorem for the martingale $M_t^{(\lambda)} = \lambda^{\frac{1}{2}} M_{\frac{t}{\lambda}}$.

Lemma 3.3. *As $\lambda \rightarrow 0$, the martingale $M_t^{(\lambda)} = \lambda^{\frac{1}{2}} M_{\frac{t}{\lambda}}$ converges in law with respect to the uniform metric to a standard Brownian motion \mathbf{B}' over the interval $t \in [0, T]$.*

Proof. To prove the central limit theorem, we prove the following:

- (i). For each $t \in \mathbb{R}^+$, the predictable quadratic variation process $\langle M^{(\lambda)} \rangle_t$ converges in probability to t as $\lambda \rightarrow 0$.
- (ii). For any $\epsilon > 0$, then as $\lambda \rightarrow 0$

$$\mathbb{P}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} (M_r - M_{r-})^2 > \frac{\epsilon}{\lambda} \right] \rightarrow 0.$$

By [29, Thm. VIII.2.13] this is sufficient to prove that $M_t^{(\lambda)}$ converges in law to a Brownian motion.

- (i) We prove a somewhat stronger statement. Note that

$$\langle M^{(\lambda)} \rangle_t - t = \lambda \int_0^{\frac{t}{\lambda}} dr (\mathcal{Q}_\lambda(P_r) - 1).$$

For the expectation of the supremum of the difference between $\langle M^{(\lambda)} \rangle_t$ and t over the interval $[0, T]$, we have

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{t \in [0, T]} \left| \langle M^{(\lambda)} \rangle_t - t \right| \right] &\leq \lambda \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr |\mathcal{Q}_\lambda(P_r) - 1| \right] \\ &\leq CT \lambda^{\frac{1}{2}} \left(\lambda^{\frac{1}{2}} + \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} \lambda^{\frac{1}{2}} |P_r| \right] + \lambda^{\frac{1}{2}} \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} \lambda |P_r|^2 \right] \right) \\ &= O(\lambda^{\frac{1}{2}}), \end{aligned}$$

where the second inequality is for some $C > 0$ by Part (4) of Proposition 3.1. The expectations in the second line above are bounded uniformly for $\lambda < 1$ by Lemma 2.2 since $|P_r| \leq 2^{\frac{1}{2}} H_r^{\frac{1}{2}}$. The above implies that $\langle M^{(\lambda)} \rangle_t$ converges in probability to t .

(ii). Recall that \mathcal{N}_t is the number of collisions over the time interval $[0, t]$ and that $t_1, \dots, t_{\mathcal{N}_t}$ are the corresponding jump-times.

$$\begin{aligned} \mathbb{P}^{(\lambda)} \left[\sup_{0 \leq r \leq \frac{T}{\lambda}} (M_r - M_{r-})^2 > \frac{\epsilon}{\lambda} \right] &\leq \frac{\lambda}{\epsilon} \mathbb{E}^{(\lambda)} \left[\left(\sum_{n=1}^{\mathcal{N}_{\frac{T}{\lambda}}} (M_{t_n} - M_{t_n^-})^4 \right)^{\frac{1}{2}} \right] \\ &\leq \frac{\lambda}{\epsilon} \mathbb{E}^{(\lambda)} \left[\sum_{n=1}^{\mathcal{N}_{\frac{T}{\lambda}}} (M_{t_n} - M_{t_n^-})^4 \right]^{\frac{1}{2}} \\ &= \frac{\lambda}{\epsilon} \mathbb{E}^{(\lambda)} \left[\sum_{n=1}^{\mathcal{N}_{\frac{T}{\lambda}}} \mathbb{E}^{(\lambda)} \left[(M_r - M_{r-})^4 \mid P_{r-}, \mathcal{N}_r = \mathcal{N}_{r-} + 1 \right] \Big|_{r=t_n} \right]^{\frac{1}{2}} \\ &= \frac{\lambda}{\epsilon} \mathbb{E}^{(\lambda)} \left[\sum_{n=1}^{\mathcal{N}_{\frac{T}{\lambda}}} \frac{\Pi_{\lambda,4}(P_{t_n^-})}{\mathcal{E}_\lambda(P_{t_n^-})} \right]^{\frac{1}{2}} = \frac{\lambda}{\epsilon} \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr \Pi_{\lambda,4}(P_r) \right]^{\frac{1}{2}}. \quad (3.3) \end{aligned}$$

The second inequality is Jensen's, and the first inequality is Chebyshev's followed by the elementary relation

$$\sup_{1 \leq m \leq n} a_m \leq \left(\sum_{m=1}^n a_m^2 \right)^{\frac{1}{2}}, \quad a_n \geq 0.$$

The second equality uses that a jump for M_r is a jump for P_t (since they differ by a continuous quantity) and that the conditional expectation for $(P_r - P_{r-})^4$ given the value P_{r-} and the information that r is a jump-time can be written in terms of $\Pi_\lambda^{(4)}(P_{r-})$ and $\mathcal{E}_\lambda(P_{r-})$ as

$$\mathbb{E}^{(\lambda)} \left[(M_r - M_{r-})^4 \mid P_{r-}, \mathcal{N}_r = \mathcal{N}_{r-} + 1 \right] = \mathbb{E}^{(\lambda)} \left[(P_r - P_{r-})^4 \mid P_{r-}, \mathcal{N}_r = \mathcal{N}_{r-} + 1 \right] = \frac{\Pi_{\lambda,4}(P_{r-})}{\mathcal{E}_\lambda(P_{r-})}.$$

The last equality follows because the jump times t_n occur with Poisson rate $\mathcal{E}_\lambda(P_r)$.

Squaring the right side of (3.3),

$$\left(\frac{\lambda}{\epsilon} \right)^2 \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr \Pi_{\lambda,4}(P_r) \right] \leq \left(\frac{\lambda}{\epsilon} \right)^2 \mathbb{E}^{(\lambda)} \left[\int_0^{\frac{T}{\lambda}} dr (1 + \lambda |P_r|)^5 \right] \leq \frac{\lambda}{\epsilon^2} \mathbb{E}^{(\lambda)} \left[\lambda \int_0^{\frac{T}{\lambda}} dr (1 + \lambda 2^{\frac{1}{2}} H_r^{\frac{1}{2}})^5 \right],$$

where we have applied Part (6) of Proposition 3.1 in the first inequality, and the bound $|P_r|^2 \leq 2H_r$ for the second. By Lemma 2.2, the expectation on the right side is uniformly bounded for $\lambda < 1$. Thus, the above is $O(\lambda)$ and the left side of (3.3) is $O(\lambda^{\frac{1}{2}})$, and we have proven the Lindberg condition. \square

4 Bounds involving the cumulative potential forcing

Section 4.1 defines the splitting structure that we use, and includes a few elementary formulas regarding it. Section 4.4 makes the estimates that convert the rapid oscillations of the force $\frac{dV}{dx}(X_r)$ when $|P_r| \gg 1$ into decay for various integrals of the force. Finally, Section 4.5 bounds the momentum drift $D_{\frac{t}{\lambda}} = \int_0^{\frac{t}{\lambda}} dr \frac{dV}{dx}(X_r)$ for $\lambda \ll 1$.

4.1 Nummelin splitting

We must start by introducing some extra artificial structure to our process. The split process that we define here is a reduced version of that in [21]. In the context of a larger probability space, the drift in momentum $D_t = \int_0^t dr \frac{dV}{dx}(X_r)$ may be viewed as a martingale plus a few small “boundary” terms. This allows us to apply martingale techniques. For those familiar with the terminology related to Nummelin splitting, we outline the extension of the process as follows: We introduce a resolvent chain embedded in the original process, we split the chain using Nummelin’s technique, and we extend the resolvent chain to a non-Markovian process which contains an embedded version of the original process.

Let (e_m) be a sequence of mean-1 exponential random variables which are independent of each other and of the process (X_t, P_t) , and let $\tau_n = \sum_{m=1}^n e_m$ with the convention $\tau_0 = 0$. The τ_n will be referred to as the *partition times*. Define \mathbf{N}_t to be the number of non-zero τ_n less than t , and the Markov chain $\sigma_n = (X_{\tau_n}, P_{\tau_n}) \in \Sigma$, which is referred to as the *resolvent chain*. The resolvent chain has the same invariant probability density as the original process. Let \mathcal{T}_λ be the transition kernel for the chain, which acts on functions from the left and measures from the right. Recall that for a Markov chain, an *atom* is a nonempty set α such that the probability transitions starting from a point $x \in \alpha$ are independent of x . An atom is said to be *recurrent* if, when starting from the atom, the probability of returning to the atom in the future is non-zero. In general, a Markov chain does not necessarily have a recurrent atom. The splitting technique we outline presently, originally due to Nummelin, allows us to create a recurrent atom from a certain minorization condition. The idea is to extend the state space Σ to $\tilde{\Sigma} = \Sigma \times \{0, 1\}$ in order to construct a chain $(\tilde{\sigma}_n) \in \tilde{\Sigma}$ with a recurrent atom and have the statistics for (σ_n) embedded in the first component of $(\tilde{\sigma}_n)$. Let ν be a probability measure on Σ and $h : \Sigma \rightarrow [0, 1)$ be such that

$$\mathcal{T}_\lambda(s_1, ds_2) \geq h(s_1)\nu(ds_2). \quad (4.1)$$

By Part (1) of Proposition 4.3, we may even pick a single pair ν and h satisfying (4.1) for all $\lambda < 1$. We have the following transition rates from the state $(s_1, z_1) \in \tilde{\Sigma}$ to the infinitesimal

region (ds_2, z_2) :

$$\tilde{\mathcal{T}}_\lambda(s_1, z_1; ds_2, z_2) = \begin{cases} \frac{1-h(s_2)}{1-h(s_1)}(\mathcal{T}_\lambda - h \otimes \nu)(s_1, ds_2) & z_1 = z_2 = 0, \\ \frac{h(s_2)}{1-h(s_1)}(\mathcal{T}_\lambda - h \otimes \nu)(s_1, ds_2) & z_1 = 1 - z_2 = 0, \\ (1 - h(s_2))\nu(ds_2) & z_1 = 1 - z_2 = 1, \\ h(s_2)\nu(ds_2) & z_1 = z_2 = 1. \end{cases}$$

Given a measure μ on Σ , we refer to its *splitting* $\tilde{\mu}$ as the measure on $\tilde{\Sigma}$ given by

$$\tilde{\mu}(ds, z) = \chi(z = 0)(1 - h(s))\mu(ds) + \chi(z = 1)h(s)\mu(ds). \quad (4.2)$$

In particular, the split chain is taken to have initial distribution given by the splitting of the initial distribution for the original (pre-split) chain. The set $\Sigma \times 1$ is an atom since the transition measure from $(s_1, 1)$ is independent of s_1 . Moreover, it is a recurrent atom, since our original process is exponentially ergodic to $\Psi_{\infty, \lambda}$ by Theorem A.1, and as a consequence, the split chain is exponentially ergodic with respect to the invariant state $\tilde{\Psi}_{\infty, \lambda}$ (see Part (2) of Proposition 4.3) which has $\tilde{\Psi}_{\infty, \lambda}(\Sigma \times 1) = \Psi_{\infty, \lambda}(h) > 0$. Notice that the conditional probability that $z_2 = 1$ given s_1, z_1, s_2 is determined by a coin with head-probability $h(s_2)$.

Using the law for the split chain $(\tilde{\sigma}_n)$ we may construct a split process $(\tilde{S}_t) \in \tilde{\Sigma}$ and a sequence of times $\tilde{\tau}_n$ with the recipe below. The $\tilde{\tau}_n$ should be thought of as the partition times τ_n embedded in the split statistics, although we temporarily denote them differently to emphasize their axiomatic role in the construction of the split process. Let $\tilde{\tau}_n$ and $\tilde{S}_t = (S_t, Z_t)$ be such that

1. $0 = \tilde{\tau}_0$, $\tilde{\tau}_n \leq \tilde{\tau}_{n+1}$, and $\tilde{\tau}_n \rightarrow \infty$ almost surely.
2. The chain $(\tilde{S}_{\tilde{\tau}_n})$ has the same law as $(\tilde{\sigma}_n)$.
3. For $t \in [\tilde{\tau}_n, \tilde{\tau}_{n+1})$, then $Z_t = Z_{\tilde{\tau}_n}$.
4. Conditioned on the information known up to time $\tilde{\tau}_n$ for \tilde{S}_t , $t \in [0, \tilde{\tau}_n]$ and $\tilde{\tau}_m$, $m \leq n$, and also the value $\tilde{S}_{\tilde{\tau}_{n+1}}$, the law for the trajectories S_t , $t \in [\tilde{\tau}_n, \tilde{\tau}_{n+1}]$ (which includes the length $\tilde{\tau}_{n+1} - \tilde{\tau}_n$) agrees with the law for the original process conditioned on knowing the values $S_{\tilde{\tau}_n}$ and $S_{\tilde{\tau}_{n+1}}$.

The marginal distribution for the first component S_t agrees with the original process and the times $\tilde{\tau}_n$ are independent mean-1 exponential random variables which are independent of S_t . Of course, the times $\tilde{\tau}_n$ are not independent of the process \tilde{S}_t , and we emphasize that the increment $\tilde{\tau}_{n+1} - \tilde{\tau}_n$ is not exponential given the state $\tilde{S}_{\tilde{\tau}_n}$. The process \tilde{S}_t is not Markovian, although, as emphasized in [21], the process $(S_t, Z_t, S_{\tau(t)}) \in \Sigma \times \{0, 1\} \times \Sigma$ is Markovian, where $\tau(t)$ is the first partition time $\tilde{\tau}_n$ following time t . We now drop the tilde from $\tilde{\tau}_n$, and use $(\tilde{\sigma}_n)$ to denote the sequence $(\tilde{S}_{\tilde{\tau}_n})$. We refer to the statistics of the split process by $\tilde{\mathbb{E}}^{(\lambda)}$ and $\tilde{\mathbb{P}}^{(\lambda)}$ for expectations and probabilities, respectively.

Now that we have defined the split process \tilde{S}_t , we can proceed to define the “life cycles”. Let R'_m be the value $\tau_{\tilde{n}_m}$ for $\tilde{n}_m = \min\{n \in \mathbb{N} \mid \sum_{k=0}^n \chi(Z_{\tau_k} = 1) = m\}$. In other words, R'_m is the m th partition time to visit the atom set $\Sigma \times 1$, and we use the convention that $R'_0 = 0$. Define R_m , $m \geq 1$ to be the partition time following R'_m . The m th life cycle is the time interval $[R_m, R_{m+1})$. Intuitively, it may at first seem more natural to define $S_{R'_m}$ as the beginning of the life cycle. However, since the distribution for R'_1 will depend on the initial distribution \tilde{S}_0 . It

is more stable to consider the beginning of the life cycle to be the partition time R_m following R'_m , which has distribution $\tilde{\nu}$ with respect to information known up to time R'_m . Although the conditional distribution for \tilde{S}_{R_m} is invariant of the value $\tilde{S}_{R'_m} \in \Sigma \times 1$, successive live cycles $[R_{n-1}, R_n)$, $[R_n, R_{n+1})$ are obviously not independent, since for instance a.s. $\lim_{t \nearrow R_n} S_t = S_{R_n}$. Let $d\mathbf{N}_t$ be the counting measure on \mathbb{R}^+ such that $\int_{(t_1, t_2]} d\mathbf{N}_r = \mathbf{N}_{t_2} - \mathbf{N}_{t_1}$ for $0 \leq t_1 < t_2$ (i.e. the number of partition times over the interval $(t_1, t_2]$). The following proposition is a consequence of [21] and lists the independence properties that we have.

Proposition 4.1.

1. The distribution for \tilde{S}_{R_n} is $\tilde{\nu}$ when conditioned on all information known up to time R'_n : $\tilde{\mathcal{F}}_{R'_n}$.
2. The sequence of trajectories $(\tilde{S}_t, d\mathbf{N}_t : t \in [R_n, R'_{n+1}])$ are i.i.d. for $n \geq 1$, and $(\tilde{S}_t, d\mathbf{N}_t : t \in [R_n, R'_{n+1}])$ is independent of $(\tilde{S}_t, d\mathbf{N}_t : t \in \mathbb{R}^+ - [R'_n, R_{n+1}])$.
3. The trajectory $(\tilde{S}_t, d\mathbf{N}_t : t \in [R_n, R_{n+1}])$ is independent of $(\tilde{S}_t, d\mathbf{N}_t : t \in \mathbb{R}^+ - [R'_n, R_{n+2}])$. In particular, $(\tilde{S}_t, d\mathbf{N}_t : t \in [R_n, R_{n+1}])$ is independent of $(\tilde{S}_t, d\mathbf{N}_t : t \in [R_m, R_{m+1}])$ for $|n - m| \geq 2$.

Unfortunately, notation multiplies when the splitting structure is invoked. For easier reference, we list the following frequently used symbols:

$\tilde{S}_t = (S_t, Z_t)$	State of the split process at time t .
$\tau_m \in \mathbb{R}^+$	m th partition time.
$\tilde{\sigma}_m = \tilde{S}_{\tau_m}$	m th state of the split chain.
$(\sigma_m, \zeta_m) = \tilde{\sigma}_m$	σ_m and ζ_m are the state and binary components, respectively, of $\tilde{\sigma}_m$.
$\mathbf{N}_t \in \mathbb{N}$	Number of partition times up to time t .
$R'_m \in \mathbb{R}^+$	m th partition time visiting the set $\Sigma \times 1$.
$R_m \in \mathbb{R}^+$	Partition time succeeding R'_m and the beginning of the m th life cycle.
$\tilde{N}_t \in \mathbb{N}$	Number of returns to the atom up to time t .
$\tilde{n}_m \in \mathbb{N}$	Number of partition times in interval $(0, R_m]$.
$\mu \rightarrow \tilde{\mu}$	The splitting of a measure μ on Σ as defined in (4.2).
\mathcal{F}_t	Information up to time t for the original process S_r and the τ_m .
$\tilde{\mathcal{F}}_t$	Information up to time t for the split process \tilde{S}_t and the τ_m .
$\tilde{\mathcal{F}}'_t$	Information for \tilde{S}_t and the τ_m before time R_{n+1} , where $R'_n \leq t < R'_{n+1}$.

There is some flexibility in the choice of ν and h in the criterion (4.1), although choosing them invariant of $\lambda > 0$ adds a little extra constraint. By Part (1) of Proposition 4.3, we can select a pair ν, h that is invariant of λ , and where both are functions of the energy. We will use the symbol ν for both the measure and the corresponding density.

Convention 4.2. We take ν and h of the form

$$h(s) = \mathbf{u} \frac{\chi(H(s) \leq l)}{U} \quad \text{and} \quad \nu(ds) = ds \frac{\chi(H(s) \leq l)}{U},$$

where $l = 1 + \sup_x V(x)$, $U > 0$ is the normalization constant of ν , and $\mathbf{u} \in (0, U)$ is from Part (1) of Proposition 4.3.

The compact support of h means that the extended state space for the split dynamics is effectively $\Sigma \times 0 \cup \{s \in \Sigma \mid H(s) \leq l\} \times 1$. Any supremum, minimum, etc. over $\tilde{\Sigma}$ refers to this contracted set.

In words, Part (3) of the proposition below states the following: given the information that time r is a partition time (i.e. $\mathbf{N}_r = \mathbf{N}_{r-} + 1$) and the information $\tilde{\mathcal{F}}_{r-}$, then the probability that particle is at the atom at time r (i.e. $Z_r = 1$) is $h(S_r)$. This holds regardless of whether the state of the process \tilde{S}_t was already sitting at the atom before time r . For the partition times $r = \tau_n$, $n \in \mathbb{N}$, knowledge of the state S_r is contained in $\tilde{\mathcal{F}}_{r-}$ almost surely, since a.s. $\lim_{v \nearrow r} S_v = S_r$. This limit will hold since the partition times and the collision times in which S_r jumps will a.s. not coincide. The proof of Part (1) of Proposition 4.3 is contained in Section 6, and the other parts of the proposition are immediate consequences from the definition of the split statistics.

Proposition 4.3.

1. *There is a constant $\mathbf{u} > 0$ such that the h and ν in Convention 4.2 satisfy $\mathcal{T}_\lambda(s, ds') \geq h(s)\nu(ds')$ for all $s, s' \in \Sigma$ and $\lambda < 1$.*

Also, the transition measures $\mathcal{T}_\lambda(s, ds')$ have densities over the domains $\{s' \in \Sigma \mid H(s') \neq H(s)\}$, which have the following bound

$$\sup_{\lambda \leq 1} \sup_{\substack{H(s) > l \\ H(s) \neq H(s')}} \frac{\mathcal{T}_\lambda(s, ds')}{ds'} < \infty.$$

2. *The invariant state of both the split chain $(\tilde{\sigma}_n)$ and the split process (\tilde{S}_t) is the splitting of the invariant state of the original process, i.e.,*

$$\tilde{\Psi}_{\infty, \lambda}(s, 0) = (1 - h(s))\Psi_{\infty, \lambda}(s) \quad \text{and} \quad \tilde{\Psi}_{\infty, \lambda}(s, 1) = h(s)\Psi_{\infty, \lambda}(s).$$

Thus the “atom” has measure $\int_\Sigma ds h(s)\Psi_{\infty, \lambda}(s)$.

3. $\tilde{\mathbb{P}}^{(\lambda)}[Z_r = 1 \mid \tilde{\mathcal{F}}_{r-}, \mathbf{N}_r - \mathbf{N}_{r-} = 1] = h(S_r)$

Besides the nearly independent behavior of the process \tilde{S}_t over the intervals $[R_m, R_{m+1})$, the payoff for introducing the splitting structure includes the closed formulas in Proposition 4.4. For Part (3) and (4) of the proposition below, $\mathfrak{R}^{(\lambda)}$ is the reduced resolvent of the backward generator \mathcal{L}_λ^*

$$\mathfrak{R}^{(\lambda)}g = \int_0^\infty dr e^{r\mathcal{L}_\lambda^*}(g),$$

which operates on $g \in L^\infty(\Sigma)$ with $\Psi_{\infty, \lambda}(g) = 0$. The reduced resolvent is well-defined since the process S_t is exponentially ergodic by Theorem A.1. As $\lambda \rightarrow 0$, the expression in Part (4) is related to the diffusion constant κ appearing in [7, Thm. 1.1].

Proposition 4.4.

1. *For $g \in L^\infty(\tilde{\Sigma})$,*

$$\tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{m=0}^{\tilde{n}_1} g(\tilde{\sigma}_n) \right] = \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{m=1}^{\tilde{n}_1+1} g(\tilde{\sigma}_m) \right] = \frac{\int_{\tilde{\Sigma}} d\tilde{s} \tilde{\Psi}_\infty^{(\lambda)}(\tilde{s}) g(\tilde{s})}{\int_\Sigma ds \Psi_\infty^{(\lambda)}(s) h(s)}.$$

In particular, if $g \in L^\infty(\Sigma)$ does not depend on the binary variable, then the numerator on the right sides above is equal to $\int_{\tilde{\Sigma}} d\tilde{s} \tilde{\Psi}_\infty^{(\lambda)}(\tilde{s}) g(\tilde{s}) = \int_\Sigma ds \Psi_{\infty, \lambda}(s) g(s)$.

2. For $g \in L^\infty(\Sigma)$,

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] = \frac{\int_\Sigma ds \Psi_{\infty, \lambda}(s) g(s)}{\int_\Sigma ds \Psi_{\infty, \lambda}(s) h(s)}.$$

3. For $g \in L^\infty(\Sigma)$ with $\Psi_{\infty, \lambda}(g) = 0$ and $s_1, s_2 \in \Sigma$,

$$\tilde{\mathbb{E}}_{\tilde{\delta}_{s_1}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] - \tilde{\mathbb{E}}_{\tilde{\delta}_{s_2}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] = (\mathfrak{R}^{(\lambda)} g)(s_1) - (\mathfrak{R}^{(\lambda)} g)(s_2),$$

where $\tilde{\delta}_s$ is the splitting of the δ measure at $s \in \Sigma$.

4. For $g \in L^\infty(\Sigma)$ with $\Psi_{\infty, \lambda}(g) = 0$,

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \int_r^{R_2} dr' g(S_{r'}) \right] = \frac{\int_\Sigma ds \Psi_{\infty, \lambda}(s) g(s) (\mathfrak{R}^{(\lambda)} g)(s)}{\int_\Sigma ds \Psi_{\infty, \lambda}(s) h(s)}.$$

Proof.

Part (1): This follows as a general fact for split chains [28] when the original chain σ_n is Harris recurrent with normalizable invariant measure $\Psi_{\infty, \lambda}$. The distribution for $\tilde{\sigma}_m$ for $m = 0$ and $m = \tilde{n}_1 + 1$ is $\tilde{\nu}$, so counting over $[0, \tilde{n}_1]$ or $[1, \tilde{n}_1 + 1]$ has the same result.

Part (2): Let $g_n^{(\lambda)} : \Sigma^2 \rightarrow \mathbb{R}$ and $\mathbf{g}^{(\lambda)} : \Sigma \rightarrow \mathbb{R}$ be defined as

$$\begin{aligned} g_n^{(\lambda)}(s, s') &:= \mathbb{E}_s^{(\lambda)} \left[\left(\int_0^{\tau_1} dr g(S_r) \right)^n \middle| s' = S_{\tau_1} \right] \\ \mathbf{g}^{(\lambda)}(s) &:= \mathbb{E}_s^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) \right]. \end{aligned}$$

Also define $\tilde{\mathbf{g}}^{(\lambda)} : \tilde{\Sigma} \rightarrow \mathbb{R}$ analogously to $\mathbf{g}^{(\lambda)}(\tilde{s})$ with $\mathbb{E}_s^{(\lambda)}$ replaced by $\tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)}$.

We have the equalities

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{\tau_n}^{\tau_{n+1}} dr g(S_r) \middle| \tilde{\sigma}_n, \tilde{\sigma}_{n+1} \right] \right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} g_1^{(\lambda)}(\sigma_n, \sigma_{n+1}) \right] \\ &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \tilde{\mathbf{g}}^{(\lambda)}(\tilde{\sigma}_n) \right] = \frac{\int_{\tilde{\Sigma}} d\tilde{s} \tilde{\Psi}_{\infty, \lambda}(\tilde{s}) \tilde{\mathbf{g}}^{(\lambda)}(\tilde{s})}{\int_\Sigma ds \Psi_{\infty, \lambda}(s) h(s)}, \end{aligned} \quad (4.3)$$

where the second equality holds, since the statistics for S_r over an interval (τ_n, τ_{n+1}) given the values $\tilde{\sigma}_n = (\sigma_n, \zeta_n)$, $\tilde{\sigma}_{n+1} = (\sigma_{n+1}, \zeta_{n+1})$ is independent of ζ_n, ζ_{n+1} and is the same for the split and the original dynamics. The third equality is from Part (1).

Starting from the right side of (4.3),

$$\begin{aligned}
\frac{\int_{\tilde{\Sigma}} d\tilde{s} \tilde{\Psi}_{\infty,\lambda}(\tilde{s}) \tilde{\mathbf{g}}^{(\lambda)}(\tilde{s})}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} &= \frac{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) \mathbf{g}^{(\lambda)}(s)}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} = \frac{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) \mathbb{E}_s^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) \right]}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} \\
&= \frac{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) \mathbb{E}_s^{(\lambda)} \left[\int_0^{\infty} dr e^{-r} g(S_r) \right]}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} \\
&= \frac{\int_0^{\infty} dr e^{-r} \left(\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) \mathbb{E}_s^{(\lambda)} [g(S_r)] \right)}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} \\
&= \frac{\int_0^{\infty} dr e^{-r} \left(\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) g(s) \right)}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)} = \frac{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) g(s)}{\int_{\Sigma} ds \Psi_{\infty,\lambda}(s) h(s)}. \tag{4.4}
\end{aligned}$$

The first equality uses that $\tilde{\Psi}_{\infty,\lambda}$ has the split form in Part (2) of Proposition 4.3, and the third uses that τ_1 is a mean-1 exponential independent of S_t in the original statistics. The fourth equality is Fubini, and the fifth is due to the stationarity of $\Psi_{\infty,\lambda}$.

Part (3): The reduced resolvent $\mathfrak{R}^{(\lambda)}$ is a pointwise limit

$$(\mathfrak{R}^{(\lambda)}g)(s) = \lim_{\gamma \searrow 0} \mathbb{E}_s^{(\lambda)} \left[\int_0^{\infty} dr e^{-r\gamma} g(S_r) \right] = \lim_{\gamma \searrow 0} \tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} \left[\int_0^{\infty} dr e^{-r\gamma} g(S_r) \right],$$

where the second equality is embedding the expectation in the split statistics. However, for $s_1, s_2 \in \Sigma$,

$$\left(\tilde{\mathbb{E}}_{\tilde{s}_1}^{(\lambda)} - \tilde{\mathbb{E}}_{\tilde{s}_2}^{(\lambda)} \right) \left[\int_0^{\infty} dr e^{-r\gamma} g(S_r) \right] = \left(\tilde{\mathbb{E}}_{\tilde{s}_1}^{(\lambda)} - \tilde{\mathbb{E}}_{\tilde{s}_2}^{(\lambda)} \right) \left[\int_0^{R_1} dr e^{-r\gamma} g(S_r) \right],$$

since the distribution for the state \tilde{S}_{R_1} is $\tilde{\nu}$ regardless of the initial measure. Using the above equalities

$$\begin{aligned}
(\mathfrak{R}^{(\lambda)}g)(s_1) - (\mathfrak{R}^{(\lambda)}g)(s_2) &= \lim_{\gamma \searrow 0} \left(\tilde{\mathbb{E}}_{\tilde{s}_1}^{(\lambda)} - \tilde{\mathbb{E}}_{\tilde{s}_2}^{(\lambda)} \right) \left[\int_0^{R_1} dr e^{-r\gamma} g(S_r) \right] \\
&= \tilde{\mathbb{E}}_{\tilde{s}_1}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] - \tilde{\mathbb{E}}_{\tilde{s}_2}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right],
\end{aligned}$$

where the limits are well-defined since the process \tilde{S}_t is positive-recurrent by A.1 and hence $\tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)}[R_1]$ is finite for all $\tilde{s} \in \tilde{\Sigma}$.

Part (4): Notice that

$$\begin{aligned}
\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \int_r^{R_2} dr' g(S_{r'}) \right] &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\mathbb{E}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \int_{\tau_n}^{\tau_{n+1}} dr g(S_r) \int_r^{R_2} dr' g(S_{r'}) \mid \tilde{\mathcal{F}}_{\tau_n^-} \right] \right] \\
&= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \mathbf{f}^{(\lambda)}(\sigma_n) \right], \tag{4.5}
\end{aligned}$$

where $\mathbf{f}^{(\lambda)} : \Sigma \rightarrow \mathbb{R}$ is defined as

$$\mathbf{f}^{(\lambda)}(s) := \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) \int_r^{R_2} dr' g(S_{r'}) \right].$$

The equality (4.5) uses the strong Markov property at the times τ_n for $n \in [0, \tilde{n}_1]$, and the split δ -distribution $\tilde{\delta}_s$ is a consequence of Part (3) of Proposition 4.3. The function $\mathbf{f}^{(\lambda)}(s)$ can be rewritten as

$$\begin{aligned} \mathbf{f}^{(\lambda)}(s) &= \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) \int_r^{\tau_1} dv g(S_v) + \left(\int_0^{\tau_1} dr g(S_r) \right) \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{\tau_1}^{R_2} dr g(S_r) \mid \tilde{\mathcal{F}}_{\tau_1}^- \right] \right] \\ &= \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) \int_r^{\tau_1} dv g(S_v) + \left(\int_0^{\tau_1} dr g(S_r) \right) (\mathfrak{R}^{(\lambda)} g)(S_{\tau_1}) + c \int_0^{\tau_1} dr g(S_r) \right] \\ &= \mathbb{E}_s^{(\lambda)} \left[\int_0^{\tau_1} dr g(S_r) (\mathfrak{R}^{(\lambda)} g)(S_r) + c \int_0^{\tau_1} dr g(S_r) \right], \end{aligned} \quad (4.6)$$

where $c \in \mathbb{R}$ is the constant such that $(\mathfrak{R}^{(\lambda)} g)(s) + c = \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{R_2} dr g(S_r) \right]$ for all $s \in \Sigma$ by Part (3). The value for c depends on g and the choice of ν, h defining the Nummeling splitting. For the second equality, we have used that

$$\tilde{\mathbb{E}}^{(\lambda)} \left[\int_{\tau_1}^{R_2} dr g(S_r) \mid \tilde{\mathcal{F}}_{\tau_1}^- \right] = \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{\tau_1}}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \right] = (\mathfrak{R}^{(\lambda)} g)(S_{\tau_1}) + c,$$

where the first equality is by Part (3) of Proposition 4.3. The third equality of (4.6) follows by replacing $\tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)}$ with $\mathbb{E}_s^{(\lambda)}$, and then using a nested conditional expectation with respect to \mathcal{F}_r for $r \leq \tau_1$:

$$\mathbb{E}^{(\lambda)} \left[\int_r^{\tau_1} dv g(S_v) + (\mathfrak{R}^{(\lambda)} g)(S_{\tau_1}) \mid \mathcal{F}_r \right] = \mathbb{E}_{S_r}^{(\lambda)} \left[\int_0^{\tau_1} dv g(S_v) + (\mathfrak{R}^{(\lambda)} g)(S_{\tau_1}) \right] = (\mathfrak{R}^{(\lambda)} g)(S_r).$$

We can then plug our expression (4.6) for $\mathbf{f}^{(\lambda)}(s)$ into (4.5) and invert the first two steps of the proof to obtain that

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) \int_r^{R_2} dr' g(S_{r'}) \right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr g(S_r) (\mathfrak{R}^{(\lambda)} g)(S_r) + c \int_0^{R_1} dr g(S_r) \right].$$

Finally, we can apply Part (2). The constant c disappears from the expression since $\Psi_{\infty, \lambda}(g) = 0$.

□

Lemma 4.5. *Let the process \tilde{M}_t be defined as*

$$\tilde{M}_t = \sum_{n=1}^{\tilde{N}_t} \left(\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) - (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_n}) + (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_{n+1}}) \right).$$

The process \tilde{M}_t is a martingale with respect to the filtration $\tilde{\mathcal{F}}'_t$. Moreover, the predictable quadratic variation $\langle \tilde{M} \rangle_t$ has the form

$$\langle \tilde{M} \rangle_t = \sum_{n=1}^{\tilde{N}_t} \bar{v}_\lambda(S_{R_n}),$$

where $\bar{v}_\lambda : \Sigma \rightarrow \mathbb{R}^+$ is defined as

$$\bar{v}_\lambda(s) = 2\tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_r) \right] + \int_\Sigma d\nu(s') \left((\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(s') \right)^2 - \left((\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(s) \right)^2.$$

In the above, $\tilde{\delta}_s$ is the splitting of the δ -distribution at s .

Proof. The jump times R'_n for the martingale \tilde{M}_t are predictable with respect to the filtration $\tilde{\mathcal{F}}_t$, although we will show that the values of the jumps still have mean zero with respect to the information known before the time of the jump. We can rewrite the martingale as

$$\tilde{M}_t = \sum_{n=1}^{\tilde{N}_t} \int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) - \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{R_n^-} \right] + \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_{n+1}}^{R_{n+2}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{R_{n+1}^-} \right],$$

since by Part (3) of Proposition 4.3

$$\tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{R_n^-} \right] = \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_n}}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) \right] = (\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(S_{R_n}) + c, \quad (4.7)$$

where the second equality is for some $c \in \mathbb{R}$ depending on g, ν, h by Part (3) of Proposition 4.11.

The jumps of \tilde{M}_t have mean zero since for $t < R'_n$, the conditional expectation of the n th jump given $\tilde{\mathcal{F}}'_t$ is

$$\begin{aligned} \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) - \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{R_n^-} \right] \right. \\ \left. + \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_{n+1}}^{R_{n+2}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{R_{n+1}^-} \right] \mid \tilde{\mathcal{F}}'_t \right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right] = 0. \end{aligned}$$

For the first equality the first two terms cancel by a nested conditional expectation, since $\tilde{\mathcal{F}}'_t \subset \tilde{\mathcal{F}}_{R_n^-}$. For the third term in the first equality, we use the strong Markov property at the time R_{n+1} and that $\tilde{S}_{R_{n+1}}$ has distribution $\tilde{\nu}$ when conditioned on $\tilde{\mathcal{F}}_{R_n}$ (and thus when conditioned on $\tilde{\mathcal{F}}'_t$) by Part (1) of Proposition 4.1. The second equality is by Part (2) of Proposition 4.4 with $g(x, p) = \frac{dV}{dx}(x)$, since $\Psi_{\infty, \lambda}(\frac{dV}{dx}) = 0$. Thus \tilde{M}_t is a martingale.

For $t \in [R'_{n-1}, R'_n)$, the σ -algebra $\tilde{\mathcal{F}}'_t$ contains all information before time R_n (i.e. $\tilde{\mathcal{F}}'_t = \tilde{\mathcal{F}}_{R_n^-}$). The predictable quadratic variation $\langle \tilde{M} \rangle_t$ must have the form of a discrete sum over $\sum_{n=1}^{\tilde{N}_t}$, since the jump times R'_n are predictable. As a consequence of Part (3) of Proposition 4.3, the conditional distribution for \tilde{S}_{R_n} given $\tilde{\mathcal{F}}'_t$ is $\tilde{\delta}_{S_{R_n}}$. The variance of the n th jump for \tilde{M}_t conditioned on $\tilde{\mathcal{F}}'_t$ is $\bar{v}_\lambda(S_{R_n})$, where

$$\begin{aligned} \bar{v}_\lambda(s) &= \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\left(\int_0^{R_1} dr \frac{dV}{dx}(X_r) - \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right] + \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_{R_1}^{R_2} dr \frac{dV}{dx}(X_r) \right] \right)^2 \right] \\ &= 2\tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \int_r^{R_2} dr' \frac{dV}{dx}(X_{r'}) \right] + \int_\Sigma d\nu(s') \left(\tilde{\mathbb{E}}_{\tilde{\delta}_{s'}}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right] \right)^2 \\ &\quad - \left(\tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right] \right)^2. \end{aligned}$$

This expression for $\bar{v}_\lambda(s)$ can be written in terms of $(\mathfrak{R}^{(\lambda)} \frac{dV}{dx})$ as in the statement of this lemma by using that $(\mathfrak{R}^{(\lambda)} \frac{dV}{dx})(s) = \tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right] - c$ and the same reasoning in the proof of Part (4) of Proposition 4.4. \square

4.2 Behavior in the $\lambda \ll 1$ regime

In the section, we are interested in a few estimates regarding “how long” the life cycles typically are for small λ and in the asymptotics transition rates \mathcal{T}_λ . We start off with a little lemma regarding the length up to the first partition time τ_1 depending on the initial state $\tilde{s} \in \tilde{S}$.

Lemma 4.6. *For $c = \frac{U}{\mathbf{u}} \vee \frac{U}{U-\mathbf{u}}$, the following inequality holds:*

$$\sup_{\lambda < 1} \sup_{\tilde{s} \in \tilde{S}} \tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} [\delta_t(\tau_1)] \leq ce^{-t},$$

where $\tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} [\delta_t(\tau_1)]$ refers to the density of the random variable τ_1 at the value $t \geq 0$. As a consequence,

$$\sup_{\lambda < 1} \sup_{\tilde{s} \in \tilde{S}} \tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} [(R_1 - R'_1)^n] < \infty.$$

Proof. In the original dynamics, τ_1 has a mean-1 exponential distribution regardless of the initial state. Splitting the distribution starting from $s \in S$ yields the equality

$$e^{-t} = \mathbb{E}_s^{(\lambda)} [\delta_t(\tau_1)] = (1 - h(s)) \tilde{\mathbb{E}}_{(s,0)}^{(\lambda)} [\delta_t(\tau_1)] + h(s) \tilde{\mathbb{E}}_{(s,1)}^{(\lambda)} [\delta_t(\tau_1)].$$

For s with $H(s) > l$, we have $h(s) = 0$ and no splitting occurs, and thus $c = 1$ is sufficient for the inequality. For s with $H(s) \leq l$, then $c = \inf_{H(s) \leq l} \frac{1}{h(s)} \vee \frac{1}{1-h(s)} = \frac{U}{\mathbf{u}} \vee \frac{U}{U-\mathbf{u}}$, where l , \mathbf{u} , and U are defined as in Convention 4.2. The bound for the moments of $R_1 - R'_1$ follows because R_1 is the first partition time after R'_1 , and by the strong Markov property for the chain $\tilde{\sigma}_n = \tilde{S}_{\tau_n}$, since $R'_1 = \tau_{\tilde{n}_1}$ for the hitting time \tilde{n}_1 . \square

For any fixed λ , the original dynamics is exponentially ergodic to the equilibrium state $\Psi_{\infty,\lambda}$. Thus the split dynamics converges exponentially to the $\tilde{\Psi}_{\infty,\lambda}$, and the time span R_1 and the number of partition times \tilde{n}_1 during a single life cycle will have finite expectation. However, the process $S_t = (X_t, P_t)$ behaves more and more like a random walk in the P_t variable as $\lambda \rightarrow 0$, so we should expect that

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [R_1] \longrightarrow \infty \quad \text{and} \quad \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1] \longrightarrow \infty,$$

as $\lambda \rightarrow 0$, since the return times of a random walk to the region around the origin have infinite expectation. However, random walks do have finite fractional moments $< \frac{1}{2}$ for their return times, and Part (2) of Proposition 4.7 states the analogous property for our process.

Proposition 4.7. *Let \tilde{n}_1 and R_1 be defined as above.*

1. *There is a $C > 0$ such that for $\lambda < 1$*

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1] \leq C\lambda^{-\frac{1}{2}} \quad \text{and} \quad \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [R_1] \leq C\lambda^{-\frac{1}{2}}.$$

2. Each fractional moment $0 < \alpha < \frac{1}{2}$ is uniformly bounded for $\lambda < 1$

$$\sup_{\lambda < 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[\tilde{n}_1^\alpha] < \infty, \quad \text{and} \quad \sup_{\lambda < 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[R_1^\alpha] < \infty.$$

Proof.

Part (1): By Part (1) of Proposition 4.4 for the constant function $g(s) = 1$

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[\tilde{n}_1 + 1] = \frac{1}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)} = \lambda^{-\frac{1}{2}} \frac{(2\pi)^{\frac{1}{2}}}{\int_{\Sigma} ds h(s)} + O(1),$$

where the order equality is for small λ . The inequality follows similarly by Part (2) of Proposition 4.4 for $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[R_1]$.

Part (2): We can prove the result through the Laplace transform by showing that there is a $C > 0$ such that

$$\sup_{\lambda < 1} |\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma \tilde{n}_1}] - 1| \leq C\gamma^{\frac{1}{2}} \quad \text{and} \quad \sup_{\lambda < 1} |\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R_1}] - 1| \leq C\gamma^{\frac{1}{2}}.$$

The proof for R_1 and \tilde{n}_1 are similar, and we focus on R_1 . Also, it is sufficient to prove the result with R'_1 rather than R_1 , since, by Lemma 4.6 $R_1 - R'_1$ has finite expectation. We will study the following regimes for γ :

- (i). $\gamma < \lambda$,
- (ii). $\lambda \leq \gamma$ and γ sufficiently small.

The case (i) can be shown with a simple linearization around $\gamma = 0$. As a result of Part (1), there exists a $C' > 0$ such that

$$|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R_1}] - 1| \leq \gamma \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[R_1] \leq C'\gamma\lambda^{-\frac{1}{2}}.$$

When $\gamma < \lambda$, the bound on the right side is smaller than $C'\gamma^{\frac{1}{2}}$.

For the regime (ii) we can no longer rely on the first derivative of the Laplace transform, since the upper bound is growing as $O(\lambda^{-\frac{1}{2}})$. In the analysis below, we will work toward showing that there is a $c > 0$ such that for all $0 \leq t \leq \lambda^{-1}$ and $\lambda \leq 1$

$$|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}] - 1| \leq \frac{1}{c\mathbb{E}_{\nu}^{(\lambda)}[\int_0^{\gamma^{-1}} dr \mathcal{A}_{\lambda}^+(S_r)] - c^{-1}}. \quad (4.8)$$

By Lemma 2.4, there is $c' > 0$ such that all $\gamma > 0$ and s in a compact set

$$\mathbb{E}_s^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr \mathcal{A}_{\lambda}^+(S_r)\right] \geq c'\gamma^{-\frac{1}{2}} - c'^{-1}.$$

With the result of (i), this is sufficient to prove the result.

We give some preliminary bounds which will be useful. The difference between $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}]$ and 1 is smaller than

$$|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}] - 1| \leq \frac{|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}] - 1|}{\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}]} \leq \left(\sum_{m=1}^{\infty} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}]^m\right)^{-1} \leq \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma R'_m}\right]^{-1}. \quad (4.9)$$

The third inequality follows, since

$$\left(\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_1}]\right)^m = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma \sum_{n=1}^m R'_n - R_{n-1}}] \geq \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R'_m}],$$

where the $R'_n - R_{n-1}$ are independent by Part (2) of Proposition 4.1 and distributed as R'_1 when the initial distribution of the split process is $\tilde{\nu}$, and the inequality is from $\sum_{n=1}^m R'_n - R_{n-1} \leq R'_m$. By (4.9), it is sufficient to give a lower bound for $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[\sum_{m=1}^{\infty} e^{-\gamma R_m}]$. This term can be rewritten as

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma R'_m}\right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma \tau_m} \chi(\zeta_m = 1)\right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma \tau_m} h(\sigma_m)\right],$$

where the first equality is from the definition of the times $R'_m = \tau_{\tilde{n}_m}$, and the second equality is by Part (3) of Proposition 4.3. The right side above is equal to

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma \tau_m} h(\sigma_m)\right] &= \mathbb{E}_{\nu}^{(\lambda)}\left[\sum_{m=1}^{\infty} e^{-\gamma \tau_m} h(\sigma_m)\right] = \mathbb{E}_{\nu}^{(\lambda)}\left[\int_0^{\infty} dr e^{-\gamma r} h(S_r)\right] \\ &\geq e^{-1} \mathbb{E}_{\nu}^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr h(S_r)\right]. \end{aligned} \quad (4.10)$$

The first equality uses that the argument of the expectation is a function of the pre-split process, and that $\tilde{\nu}$ is the splitting on ν . The last inequality is an equality for the first term, since the terms $e^{-\gamma \tau_m} h(\sigma_m) = e^{-\gamma \tau_m} h(S_{\tau_m})$ in the sum occur with Poisson rate one, and thus $\sum_{m=1}^{N_t} e^{-\gamma \tau_m} h(S_{\tau_m}) - \int_0^t dr e^{-\gamma r} h(S_r)$ is a martingale.

So far, we have shown that

$$|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[e^{-\gamma R_1}] - 1| \leq \frac{1}{e^{-1} \mathbb{E}_{\nu}^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr h(S_r)\right]}. \quad (4.11)$$

Now, we find a lower bound $\mathbb{E}_{\tilde{\nu}}^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr h(S_r)\right]$ in terms of the same expression except with h replaced by \mathcal{A}_{λ}^+ . Define the constant

$$u_{\lambda} = \frac{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) \mathcal{A}_{\lambda}^+(s)}.$$

By the triangle inequality and going to the split statistics,

$$\begin{aligned} &\left|\mathbb{E}_{\nu}^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr h(S_r)\right] - u_{\lambda} \mathbb{E}_{\nu}^{(\lambda)}\left[\int_0^{\gamma^{-1}} dr \mathcal{A}_{\lambda}^+(S_r)\right]\right| \\ &\leq \left|\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{n=1}^{\tilde{N}_{\gamma^{-1}-1}} \int_{R_{n-1}}^{R_n} dr \left(h(S_r) - u_{\lambda} \mathcal{A}_{\lambda}^+(S_r)\right)\right]\right| + \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\int_{R_n}^{R_{n+1}} dr \left(h(S_r) + u_{\lambda} \mathcal{A}_{\lambda}^+(S_r)\right)\right] \\ &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[\tilde{N}_{\gamma^{-1}} - 1] \left|\frac{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) \left(h(s) - u_{\lambda} \mathcal{A}_{\lambda}^+(s)\right)}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)}\right| + \frac{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) \left(h(s) + u_{\lambda} \mathcal{A}_{\lambda}^+(s)\right)}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)} = 2, \end{aligned}$$

where \tilde{N}_t is the number of life cycles to be completed by time t . The inequality covers the leftover interval $[R_n, \gamma^{-1}]$ of the integration. For the equality, we have used the Markov property

at the times R_n (since the embedded chain $\sigma_n = S_{R_n}$ is a Markov chain) and Part (2) of Proposition 4.4. Moreover,

$$u_\lambda = \frac{\int_\Sigma ds e^{-\lambda H(s)} h(s)}{\int_\Sigma ds e^{-\lambda H(s)} \mathcal{A}_\lambda^+(s)} = \frac{\int_\Sigma ds h(s)}{\int_\Sigma ds \mathcal{A}_\lambda^+(s)} + O(\lambda) = \mathbf{u} + O(\lambda) \geq \frac{\mathbf{u}}{2},$$

where $\mathbf{u} = \int_\Sigma ds h(s)$, and the inequality is for λ small enough. The third equality is by Part (5) of Proposition 4.15. These observation imply that

$$\mathbb{E}_\nu^{(\lambda)} \left[\int_0^{\gamma^{-1}} dr h(S_r) \right] \geq \frac{\mathbf{u}}{2} \mathbb{E}_\nu^{(\lambda)} \left[\int_0^{\gamma^{-1}} dr \mathcal{A}_\lambda^+(S_r) \right] - 2.$$

Plugging this inequality into (4.11) gives (4.8). □

The following lemma shows that the transition rates given by $\tilde{\mathcal{T}}_\lambda$ and \mathcal{T}_λ are close to those for $\tilde{\mathcal{T}}_0$ and \mathcal{T}_0 for small λ provided that the momenta for the starting point is small enough and the size of the jump in momentum is not too large. The second part of Lemma 4.8 could have a bound of order $\lambda^{\frac{3}{4}}$ with more work. The constants \mathbf{u}, U in the statement of Part (1) are from Convention 4.2.

Lemma 4.8. *Let $|p| \leq \lambda^{-\frac{1}{4}}$ and $|p - p'| \leq \lambda^{-\frac{1}{9}}$.*

1. *There exists $\alpha, c > 0$ such that for small enough λ*

$$|\mathcal{T}_\lambda(x, p; dx', dp') - \mathcal{T}_0(x, p; dx', dp')| \leq c\lambda^{\frac{1}{4}} \mathcal{T}_0(x, p; dx', dp').$$

Consequently, for $b = \frac{U - \mathbf{u}}{\mathbf{u}}$

$$|\tilde{\mathcal{T}}_\lambda(x, p, z; dx', dp', z') - \tilde{\mathcal{T}}_0(x, p, z; dx', dp', z')| \leq cb\lambda^{\frac{1}{4}} \tilde{\mathcal{T}}_0(x, p, z; dx', dp', z').$$

2. *For $F \in L^\infty(\tilde{\Sigma})$, define $g_v^{(\lambda)}(s, s') = \mathbb{E}_s^{(\lambda)} \left[\left(\int_0^\tau dr F(X_r) \right)^v \mid S_\tau = s' \right]$. For each $v \geq 1$, there exists $c_v > 0$ such that for all $\lambda < 1$*

$$\int_\Sigma dx' dp' \mathcal{T}_0(x, p; dx', dp') \left| g_v^{(\lambda)}(x, p; x', p') - g_v^{(0)}(x, p; x', p') \right| \leq c_v \lambda^{\frac{1}{2}}.$$

Proof.

Part (1): The inequality for the difference between $\tilde{\mathcal{T}}_\lambda(s, z; s', z')$ and $\tilde{\mathcal{T}}_0(s, z; s', z')$ follows from the inequality for the difference of \mathcal{T}_λ and \mathcal{T}_0 by an analysis of the cases $z \in \{0, 1\}$. The difference is identically zero when $z = 1$. For $z = 0$,

$$\begin{aligned} |\tilde{\mathcal{T}}_\lambda(s, 0; ds', z') - \tilde{\mathcal{T}}_0(s, 0; ds', z')| &= \mathbf{h}(s', z') |\mathcal{T}_\lambda(s, ds') - \mathcal{T}_\lambda(s, ds')| \\ &\leq \mathbf{h}(s', z') c\lambda^{\frac{1}{4}} \mathcal{T}_0(s, ds') \leq (1 - h(s))^{-1} c\lambda^{\frac{1}{4}} \tilde{\mathcal{T}}_0(s, 0; ds', 0), \end{aligned}$$

where $\mathbf{h}(s, 1) = 1 - \mathbf{h}(s, 0) = h(s)$. Moreover, $(1 - h(s))^{-1} \leq \frac{U - \mathbf{u}}{\mathbf{u}} = b$ by our choice of h in Convention 4.2. The second inequality above follows by discarding the second term on right side of the equality

$$(1 - h(s')) \mathcal{T}_0(s, ds') = (1 - h(s)) \tilde{\mathcal{T}}_0(s, 0; ds', 0) + (1 - h(s')) h(s) d\nu(s).$$

Now we focus on the difference of \mathcal{T}_λ and \mathcal{T}_0 . Let $\Gamma_t^{(n)} = \mathbb{R}^n \times S_n(t)$ and $\Gamma_t = \cup_{n=0}^\infty \Gamma_t^{(n)}$ for n -simplicies $S_n(t) = \{(t_1, \dots, t_n) \in [0, t]^n \mid t_m \leq t_{m+1}\}$. Given an element $\xi = (p_1, t_1; \dots; p_n, t_n) \in \Gamma_t$ let $(x_r(\xi), p_r(\xi))$ be the trajectory which begins from (x, p) and evolves deterministically according to the Hamiltonian H over the intervals $[t_m, t_{m+1})$ (with the convention $t_0 = 0$ and $t_{n+1} = t$), and with $p_{t_m}(\xi) = p_m$ for $m \in [1, n]$. In other words, there are collisions at the times t_m at which the momentum of the particle jumps to p_m . The kernel \mathcal{T}_λ can be written as

$$\mathcal{T}_\lambda(x, p; dx', dp') = \mathbb{E} \left[\int_{\Gamma_\tau} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) D_\tau^{(\lambda)}(\xi) \right],$$

where the expectation is with respect to the exponential random variable τ , and the density $D_\tau^{(\lambda)}(\xi)$ has the form

$$D_\tau^{(\lambda)}(\xi) = e^{-\int_{t_n}^\tau dr \mathcal{E}_\lambda(p_r(\xi))} \mathcal{J}_\lambda(p_{t_n^-}, p_{t_n}) e^{-\int_{t_{n-1}}^{t_n} dr \mathcal{E}_\lambda(p_r(\xi))} \dots \mathcal{J}_\lambda(p_{t_1^-}, p_{t_1}) e^{-\int_0^{t_1} dr \mathcal{E}_\lambda(p_r(\xi))}.$$

Let $\Gamma_{\tau, \lambda}$ be the set of all trajectories such that $|p_r| \leq 2\lambda^{-\frac{1}{4}}$ for all $r \in [0, \tau]$. By the triangle inequality,

$$\begin{aligned} & |\mathcal{T}_\lambda(x, p; dx', dp') - \mathcal{T}_0(x, p; dx', dp')| \\ & \leq \mathbb{E} \left[\int_{\Gamma_{\tau, \lambda}} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)| \right] \\ & + 2 \sup_{\lambda < 1} \mathbb{E} \left[\int_{\Gamma_{\tau, \lambda}^c} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) D_\tau^{(\lambda)}(\xi) \right]. \end{aligned} \quad (4.12)$$

For the second term in (4.12),

$$\begin{aligned} & \sup_{\lambda < 1} \mathbb{E} \left[\int_{\Gamma_{\lambda, \tau}^c} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) D_\tau^{(\lambda)}(\xi) \right] = \sup_{\lambda < 1} \mathbb{E}^{(\lambda)} [\mathcal{T}_\lambda(x_\varsigma, p_\varsigma; dx', dp')] \\ & \leq \left(\sup_{\lambda < 1} \sup_{\substack{|p|^2 > l \\ H(x, p) \neq H(x', p')}} \frac{\mathcal{T}_\lambda(x, p; dx', dp')}{dx' dp'} \right) \sup_{\lambda < 1} \mathbb{P}_{(x, p)}^{(\lambda)} \left[\sup_{0 \leq r \leq \tau} |P_r| > 2\lambda^{-\frac{1}{4}} \right] dx' dp' \\ & = O(e^{-\delta\lambda^{-\frac{1}{4}}}) dx' dp', \end{aligned} \quad (4.13)$$

where ς is the hitting time that $|P_r|$ reaches above $2\lambda^{-\frac{1}{4}}$, the constant $\delta > 0$ is picked below, $l > 0$ is from Convention 4.2, and $O(\cdot)$ refers to the limit $\lambda \rightarrow 0$. The first equality holds by the strong Markov property starting from ς and by the definition of the transition kernel \mathcal{T}_λ . The inequality is for λ small enough so that $\ell \leq 4\lambda^{-\frac{1}{2}}$, and the first term on the second line is finite by Part (1) of Proposition 4.3. The order equality is argued below and follows since τ is exponentially distributed, and the momentum jumps have uniformly bounded Gaussian tails from points $|p| \leq \lambda^{-\frac{1}{4}}$ and occur with Poisson rate $\leq \mathcal{E}_\lambda(\lambda^{-\frac{1}{4}}) \approx \frac{1}{8}$. The right term on the middle line of (4.13) is bounded by

$$\begin{aligned} \mathbb{P}_{(x, p)}^{(\lambda)} \left[\sup_{0 \leq r \leq \tau} |P_r| > 2\lambda^{-\frac{1}{4}} \right] & \leq \mathbb{P}_{(x, p)}^{(\lambda)} \left[\sum_{n=1}^{\mathcal{N}_{\varsigma \wedge \tau} - 1} |P_{t_n} - P_{t_n^-}| > \lambda^{-\frac{1}{4}} \right] \\ & \leq \mathbb{P}[\kappa_\tau \geq \lambda^{-\frac{1}{4}}] \leq e^{-\delta\lambda^{-\frac{1}{4}}} \mathbb{E}[e^{\delta\kappa_\tau}], \end{aligned} \quad (4.14)$$

where κ_t is an increasing Levy process starting at zero and with jump rates

$$\mathbf{j}(v) = \sup_{\lambda \leq 1} \sup_{\substack{|p'| \leq 2\lambda^{-\frac{1}{4}} \\ v = |p'' - p'|}} \mathcal{J}_\lambda(p', p'') \leq \frac{\eta}{2} v e^{-\frac{1}{8}v^2} \leq c e^{-\frac{1}{9}v^2}.$$

The bound on $\mathbf{j}(v)$ is for some $c > 0$. The first inequality in (4.14) holds since $|p| \leq \lambda^{-\frac{1}{4}}$ and so the total variation in the process must reach $\lambda^{-\frac{1}{4}}$ for $|P_r|$ to reach above $2\lambda^{-\frac{1}{4}}$. The second inequality in (4.14) holds since the rate of jumps $|P_r - P_{r-}| = v$ occurs with rate smaller $\mathbf{j}(v)$ at all $r \leq \varsigma \wedge \tau$, and the third inequality is Jensen's. The exponential moment on the right side of (4.14) is finite for small enough δ , since

$$\mathbb{E}[e^{\delta \kappa_\tau}] = \mathbb{E}\left[e^{\tau \int_{\mathbb{R}^+} dv \mathbf{j}(v) (e^{\delta v} - 1)}\right] = \left(\int_{\mathbb{R}^+} dv \mathbf{j}(v) (e^{\delta v} - 1) - 1\right)^{-1}.$$

We have given an upper bound of $O(e^{-\delta \lambda^{-\frac{1}{4}}})$ for (4.13). However, there is a $c' > 0$ such that

$$c' \mathcal{T}_0(x, p; dx', dp') > \lambda^{-\frac{1}{9}} e^{-\frac{1}{8}\lambda^{-\frac{2}{9}}} dx' dp' > e^{-\delta \lambda^{-\frac{1}{4}}} dx' dp', \quad (4.15)$$

where both inequalities are for λ small enough. The first inequality follows by an analysis of the contribution to \mathcal{T}_0 from the event that two collisions occurred over the time interval $[0, \tau]$ as in the proof of Part (1) of Proposition 4.3. This gives us the bound for the second term on the right side of (4.12).

Now we bound the first term on the right side of (4.12). For $\xi \in \Gamma_\tau^{(n)}$ and $m \in [0, n]$, define

$$\begin{aligned} D_{\tau, m}^{(\lambda)}(\xi) &= e^{-\int_{t_n}^\tau dr \mathcal{E}_\lambda(p_r(\xi))} \prod_{j=n-m+1}^n \mathcal{J}_\lambda(p_{t_j^-}(\xi), p_{t_j}(\xi)) e^{-\int_{t_{j-1}}^{t_j} dr \mathcal{E}_\lambda(p_r(\xi))} \\ &\quad \cdot \prod_{j=1}^{n-m} \mathcal{J}_0(p_{t_j^-}(\xi), p_{t_j}(\xi)) e^{-\int_{t_{j-1}}^{t_j} dr \mathcal{E}_0(p_r(\xi))}. \end{aligned}$$

By convention, we also set $D_{\tau, n+1}^{(\lambda)}(\xi) = D_\tau^{(\lambda)}(\xi)$ and $D_{\tau, 0}^{(\lambda)}(\xi) = D_\tau^{(0)}(\xi)$. By a telescoping sum for each n and the triangle inequality

$$\begin{aligned} &\mathbb{E}\left[\int_{\Gamma_{\tau, \lambda}} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)|\right] \\ &\leq \sum_{n=0}^{\infty} \sum_{m=0}^n \mathbb{E}\left[\int_{\Gamma_{\tau, \lambda}^{(n)}} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) |D_{\tau, m+1}^{(\lambda)}(\xi) - D_{\tau, m}^{(\lambda)}(\xi)|\right] \\ &\leq \sum_{n=0}^{\infty} n e^{nC\lambda^{\frac{1}{2}}} \mathbb{E}\left[\int_{\Gamma_{\tau, \lambda}^{(n)}} d\xi \delta(x' - x_\tau(\xi)) \delta(p' - p_\tau(\xi)) D_\tau^{(0)}(\xi)\right] \\ &\leq C\lambda^{\frac{1}{2}} \mathbb{E}_{(x, p)}^{(0)}\left[\delta(x' - X_\tau) \delta(p' - P_\tau) \mathcal{N}_\tau e^{C\lambda^{\frac{1}{2}} \mathcal{N}_\tau}\right], \end{aligned} \quad (4.16)$$

where we will explain the second inequality below. The third inequality follows by extending the integration over the domains $\Gamma_{\tau, \lambda}^{(n)}$ to the larger domains $\Gamma_\tau^{(n)}$.

Note that when $|p|, |p_r| \leq 2\lambda^{-\frac{1}{4}}$ for $r \in [0, t]$ and $|p - p'| \leq \lambda^{-\frac{1}{9}}$,

$$\begin{aligned} & \left| \mathcal{J}_\lambda(p, p') e^{-\int_0^t dr \mathcal{E}_\lambda(p_r)} - \mathcal{J}_0(p, p') e^{-\int_0^t dr \mathcal{E}_0(p_r)} \right| \\ & \leq \left| \mathcal{J}_\lambda(p, p') - \mathcal{J}_0(p, p') \right| e^{-\int_0^t dr \mathcal{E}_\lambda(p_r)} + \left| \mathcal{J}_0(p, p') \right| e^{-\int_0^t dr \mathcal{E}_\lambda(p_r)} - e^{-\int_0^t dr \mathcal{E}_0(p_r)} \\ & \leq (\text{const}) \left(\lambda^{\frac{3}{4}} |p - p'| + \lambda^{\frac{3}{4}} \right) \mathcal{J}_0(p, p') e^{-\frac{t}{8}} \leq C \lambda^{\frac{1}{2}} \mathcal{J}_0(p, p') e^{-\frac{t}{8}}, \end{aligned} \quad (4.17)$$

for some constant $C > 0$. Thus, for $\xi \in \Gamma_{\tau, \lambda}^{(n)}$ and $m \in [0, n]$,

$$\left| D_{\tau, m+1}^{(\lambda)}(\xi) - D_{\tau, m}^{(\lambda)}(\xi) \right| \leq C \lambda^{\frac{1}{2}} D_{\tau, m}^{(\lambda)}(\xi).$$

Thus, by an inductive argument,

$$\left| D_{\tau, m+1}^{(\lambda)}(\xi) - D_{\tau, m}^{(\lambda)}(\xi) \right| \leq C \lambda^{\frac{1}{2}} e^{m C \lambda^{\frac{1}{2}}} D_{\tau}^{(0)}(\xi) \leq C \lambda^{\frac{1}{2}} e^{n C \lambda^{\frac{1}{2}}} D_{\tau}^{(0)}(\xi).$$

Continuing with the right side of (4.16),

$$\begin{aligned} & \lambda^{\frac{1}{2}} \mathbb{E}_{(x, p)}^{(0)} \left[\delta(x' - X_\tau) \delta(p' - P_\tau) \mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \right] \\ & \leq \lambda^{\frac{1}{4}} e^{C \lambda^{\frac{1}{4}}} \mathbb{E}_{(x, p)}^{(0)} \left[\delta(x' - X_\tau) \delta(p' - P_\tau) \right] \\ & \quad + \lambda^{\frac{1}{2}} \mathbb{E}_{(x, p)}^{(0)} \left[\delta(x' - X_\tau) \delta(p' - P_\tau) \mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \chi(\lambda^{\frac{1}{4}} \mathcal{N}_\tau > 1) \right] \\ & \leq 2 \lambda^{\frac{1}{4}} \mathcal{T}_0(x, p; dx', dp') + \lambda^{\frac{1}{2}} \left(\sup_{v \in \mathbb{R}} j(v) \right) \mathbb{E}_{(x, p)}^{(0)} \left[\mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \chi(\lambda^{\frac{1}{4}} \mathcal{N}_\tau > 1) \right] dx' dp' \\ & \leq 2 \lambda^{\frac{1}{4}} \mathcal{T}_0(x, p; dx', dp') + \lambda^{\frac{1}{2}} (\text{const}) e^{-\lambda^{-\frac{1}{4}}} dx' dp', \end{aligned}$$

where $j(p' - p) = \mathcal{J}_0(p, p')$. For the first term, the second inequality is for λ small enough such that $e^{C \lambda^{\frac{1}{4}}} \leq 2$ and by the definition of \mathcal{T}_0 . The inequalities for the second term follow by

$$\begin{aligned} & \mathbb{E}_{(x, p)}^{(0)} \left[\delta(x' - X_\tau) \delta(p' - P_\tau) \mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \chi(\lambda^{\frac{1}{4}} \mathcal{N}_\tau > 1) \right] \\ & \leq \mathbb{E}_{(x, p)}^{(0)} \left[\mathbb{E} \left[\delta(x' - X_\tau) \delta(p' - P_\tau) \mid \mathcal{N}_\tau, \mathcal{F}_{t_{\mathcal{N}_\tau}^-} \right] \mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \chi(\lambda^{\frac{1}{4}} \mathcal{N}_\tau > 1) \right] \\ & \leq (\text{const}) \mathbb{E} \left[\mathcal{N}_\tau e^{C \lambda^{\frac{1}{2}} \mathcal{N}_\tau} \chi(\lambda^{\frac{1}{4}} \mathcal{N}_\tau > 1) \right] dx' dp' = (\text{const}) \sum_{n \geq \lambda^{-\frac{1}{4}}} \frac{e^{C \lambda^{\frac{1}{2}} n}}{9^n} < (\text{const}) e^{-\lambda^{-\frac{1}{4}}} dx' dp', \end{aligned}$$

where the last inequality is for λ small enough, and the equality holds since τ is mean-1 exponential and \mathcal{N}_t is a Poisson process with rate $\frac{1}{9}$. Again by the remark following (4.15), this term is smaller than a constant multiple of $\mathcal{T}_\lambda(x, p; dx', dp')$. This completes the proof.

Part (2): We can rewrite the expression

$$\begin{aligned} & \int_{\Sigma} \mathcal{T}_0(s, ds') \left| g_n^{(\lambda)}(s; s') - g_n^{(0)}(s; s') \right| \\ & = \int_{\Sigma} \mathcal{T}_0(s, ds') \left| \frac{\mathbb{E}_s^{(\lambda)} \left[\left(\int_0^\tau dr F(S_r) \right)^v \delta(S_\tau - s') \right]}{\mathcal{T}_\lambda(s, ds')} - \frac{\mathbb{E}_s^{(0)} \left[\left(\int_0^\tau dr F(S_r) \right)^v \delta(S_\tau - s') \right]}{\mathcal{T}_0(s, ds')} \right|. \end{aligned}$$

By adding and subtracting inserting the term

$$\frac{\mathbb{E}_s^{(\lambda)} \left[\left(\int_0^\tau dr \frac{dV}{dx}(X_r) \right)^v \delta(S_\tau - s') \right]}{\mathcal{T}_0(s, ds')}$$

in the absolute value and applying the triangle inequality, we attain the terms on the left sides of (4.18) and (4.19) below.

Let Γ_τ , $\Gamma_{\tau,\lambda}$, $D_\tau^{(\lambda)}$, and $x_\tau(\xi), p_\tau(\xi)$ be defined as in Part (1). We have the inequalities

$$\begin{aligned} & \int_\Sigma \left| \mathbb{E}_s^{(\lambda)} \left[\left(\int_0^\tau dr F(S_r) \right)^v \delta(S_\tau - s') \right] - \mathbb{E}_s^{(0)} \left[\left(\int_0^\tau dr F(S_r) \right)^v \delta(S_\tau - s') \right] \right| \\ & \leq \mathbb{E} \left[\int_{\Gamma_\tau} d\xi |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)| \left| \int_0^\tau dr F(x_r(\xi), p_r(\xi)) \right|^v \right] \\ & \leq \mathbb{E} \left[\int_{\Gamma_{\tau,\lambda}} d\xi |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)| \left| \int_0^\tau dr F(x_r(\xi), p_r(\xi)) \right|^v \right] + O(e^{-\delta\lambda^{-\frac{1}{4}}}) \\ & \leq 2\|F\|_\infty^v \mathbb{E} \left[\tau^v \int_{\Gamma_{\tau,\lambda}} d\xi |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)| \right], \end{aligned} \quad (4.18)$$

where the error from replacing $\Gamma_{\tau,\lambda}$ by Γ_τ is of order $O(e^{-\delta\lambda^{-\frac{1}{4}}})$ by a similar analysis as in the proof of Part (1). The factor of 2 on the last line was introduced to cover the error. Inserting a telescoping sum of $D_{\tau,n}^{(\lambda)}$, and applying the triangle inequality as in the proof of Part (1) gives

$$\begin{aligned} & \mathbb{E} \left[\tau^v \int_{\Gamma_{\tau,\lambda}} d\xi |D_\tau^{(\lambda)}(\xi) - D_\tau^{(0)}(\xi)| \right] \leq \sum_{n=0}^\infty \sum_{m=0}^n \mathbb{E} \left[\tau^v \int_{\Gamma_{\tau,\lambda}^{(n)}} d\xi |D_{\tau,m+1}^{(\lambda)}(\xi) - D_{\tau,m}^{(\lambda)}(\xi)| \right] \\ & \leq (\text{const}) \lambda^{\frac{3}{4}} \sum_{n=0}^\infty \sum_{m=0}^n \mathbb{E} \left[\tau^v \int_{\Gamma_{\tau,\lambda}^{(n)}} d\xi (|p_{t_m}(\xi) - p_{t_m^-}(\xi)| + 1) D_{\tau,m}^{(\lambda)}(\xi) \right] \\ & \leq (\text{const}) \lambda^{\frac{3}{4}} \sum_{n=0}^\infty e^{C\lambda^{\frac{1}{2}}n} \sum_{m=0}^n \mathbb{E} \left[\tau^v \int_{\Gamma_\tau^{(n)}} d\xi (|p_{t_m}(\xi) - p_{t_m^-}(\xi)| + 1) D_\tau^{(0)}(\xi) \right] \\ & = (\text{const}) \lambda^{\frac{3}{4}} \mathbb{E}^{(0)} \left[\tau^v \sum_{n=1}^{\mathcal{N}_\tau} (|P_{t_n} - P_{t_n^-}| + 1) \right] = (\text{const}) \lambda^{\frac{3}{4}} \int_{\mathbb{R}} dp' |p'| j(p'), \end{aligned}$$

For the second inequality above, we used the second inequality from (4.17). The third inequality follows from the relation $D_{\tau,m}^{(\lambda)}(\xi) \leq e^{C\lambda^{\frac{1}{2}}m} D_\tau^{(0)}(\xi)$, and then by extending the domains from $\Gamma_{\tau,\lambda}^{(n)}$ to $\Gamma_\tau^{(n)}$. Thus, we have shown that (4.18) is $O(\lambda^{\frac{3}{4}})$.

The other term is

$$\begin{aligned} & \int_\Sigma |\mathcal{T}_\lambda(s, ds') - \mathcal{T}_0(s, ds')| \left| \frac{\mathbb{E}_s^{(\lambda)} \left[\left(\int_0^\tau dr F(S_r) \right)^v \delta(S_\tau - s') \right]}{\mathcal{T}_0(s, ds')} \right| \\ & \leq \|F\|_\infty^v \int_\Sigma |\mathcal{T}_\lambda(s, ds') - \mathcal{T}_0(s, ds')| \left| \frac{\mathbb{E}_s^{(\lambda)} \left[\tau^v \delta(S_\tau - s') \right]}{\mathcal{T}_0(s, ds')} \right|. \end{aligned} \quad (4.19)$$

We split the right side into two more terms based on the events $\tau > \lambda^{-\frac{1}{4v}}$ and $\tau \leq \lambda^{-\frac{1}{4v}}$ and apply the triangle inequality

$$\begin{aligned} & \int_{\Sigma} |\mathcal{T}_{\lambda}(s, ds') - \mathcal{T}_0(s, ds')| \frac{\mathbb{E}_s^{(\lambda)}[\tau^v \delta(S_{\tau} - s')]}{\mathcal{T}_0(s, ds')} \\ & \leq \lambda^{-\frac{1}{4}} \int_{\Sigma} |\mathcal{T}_{\lambda}(s, ds') - \mathcal{T}_0(s, ds')| \frac{\mathbb{E}_s^{(\lambda)}[\delta(S_{\tau} - s')]}{\mathcal{T}_0(s, ds')} + c\lambda^{\frac{1}{4}} \int_{\Sigma} \mathbb{E}_s^{(\lambda)}[\tau^v \delta(S_{\tau} - s') \chi(\tau > \lambda^{-\frac{1}{4v}})] \\ & \leq \lambda^{-\frac{1}{4}} \int_{\Sigma} |\mathcal{T}_{\lambda}(s, ds') - \mathcal{T}_0(s, ds')| + c\lambda^{\frac{1}{4}} v! e^{-\lambda^{-\frac{1}{4v}}} = O(\lambda^{\frac{1}{2}}), \end{aligned}$$

where we have used that $\mathbb{E}_s^{(\lambda)}[\delta(S_{\tau} - s')] = \mathcal{T}_0(s, ds')$ for the first term and Part (1) for the second term. The second term on the second line is equal to $c\lambda^{\frac{1}{4}} \mathbb{E}[\tau^v \chi(\tau > \lambda^{-\frac{1}{4v}})]$ which can be computed explicitly, since τ is exponential. The integral $\int_{\Sigma} |\mathcal{T}_{\lambda}(s, ds') - \mathcal{T}_0(s, ds')|$ is $O(\lambda^{\frac{3}{4}})$ by a similar but simpler analysis than applied above for (4.18). We could have modified this argument to show that (4.19) is $O(\lambda^{\gamma})$ for any $\frac{1}{2} \leq \gamma < \frac{3}{4}$ by taking the smaller cut-off $\tau > \lambda^{\frac{1}{(\frac{3}{4}-\gamma)v}}$ in the analysis above. \square

4.3 Summing a functional over a life cycle

Let $U^{(\lambda)} : L^{\infty}(\Sigma) \rightarrow L^{\infty}(\Sigma)$ be defined as

$$(U^{(\lambda)}g)(s) = \mathbb{E}_s^{(\lambda)} \left[\int_0^{\infty} dt g(S_t) e^{-\int_0^t dr h(S_r)} \right], \quad s \in \Sigma, \quad (4.20)$$

where $h : \Sigma \rightarrow \mathbb{R}^+$ is defined as in Convention 4.2. Operators of the form (4.20) were introduced in [27]. In the proof of Lemma 4.10 below, we will need the following result, which is taken from [8].

Theorem 4.9. *Suppose that $g \in L^{\infty}(\Sigma)$ and $g \geq 0$. There is a $c > 0$ such that for any g and $|p| \leq \lambda^{-1}$*

$$\begin{aligned} (U^{(\lambda)}g)(x, p) & \leq c\|g\|_{\infty} + c|p| \sup_{H' > \frac{1}{2}\lambda^{-2}} g(x', p') + c \int_{H' \leq \frac{1}{2}\lambda^{-2}} dp' dx' (1 + |p'| \wedge |p|) g(x', p'), \\ \|U^{(\lambda)}g\|_{\infty} & \leq c\lambda^{-1} \sup_{H' > \frac{1}{2}\lambda^{-2}} g(x', p') + c \sup_{H' \leq \frac{1}{2}\lambda^{-2}} (U^{(\lambda)}g)(x', p'), \end{aligned}$$

where $H' = H(x', p')$.

Note that $U^{(\lambda)}g$ can be written in the form

$$\begin{aligned} (U^{(\lambda)}g)(s) & = \mathbb{E}_s^{(\lambda)} \left[\sum_{n=1}^{\infty} (1 - h(S_{\tau_1})) \cdots (1 - h(S_{\tau_{n-1}})) g(S_{\tau_n}) \right] \\ & = \tilde{\mathbb{E}}_{\delta_s}^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} g(S_{\tau_n}) \right] + h(s) \tilde{\mathbb{E}}_{\nu}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} g(S_{\tau_n}) \right] \geq \tilde{\mathbb{E}}_{\delta_s}^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} g(S_{\tau_n}) \right], \end{aligned} \quad (4.21)$$

where the first equality is from the proof of [17, Prop. 3.4]. The second equality embeds the quantities in the split statistics and uses Part (3) of Proposition 4.3.

The following lemma states that an additive functional of the resolvent chain $\sum_n g_\lambda(\sigma_n)$, where the summation is over a single life cycle and for $g_\lambda \geq 0$ with sufficient decay at large momentum, has arbitrary finite moments. In other words, not much happens over a single life cycle.

Lemma 4.10. *Let $g_\lambda : \Sigma \rightarrow \mathbb{R}^+$ and suppose that there is a $C > 0$ such that*

$$g_\lambda(x, p) \leq C \frac{1}{1 + |p|^2} \vee \lambda.$$

for all $\lambda < 1$. Then,

$$\sup_{\lambda < 1} \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\left(\sum_{n=0}^{\tilde{n}_1} g_\lambda(S_{\tau_n}) \right)^m \right] < \infty, \quad m \in \mathbb{N}.$$

Proof. For the case $m = 1$, we have a closed expression $f = g_\lambda$

$$\tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} f(S_{\tau_n}) \right] = \frac{\int_\Sigma ds \Psi_{\infty, \lambda}(s) f(s)}{\int_\Sigma ds \Psi_{\infty, \lambda}(s) h(s)} \quad (4.22)$$

$$\leq \frac{\int_{|p| \leq \lambda^{-1}} dx dp f(x, p) + \frac{2}{\lambda^{\frac{1}{2}}} \operatorname{erfc}(\lambda^{-\frac{1}{2}}) \sup_{|p| > \lambda^{-1}} f(x, p)}{\int_\Sigma dx dp e^{-\lambda H(x, p)} h(x, p)}, \quad (4.23)$$

where $\operatorname{erfc}(q) = \int_q^\infty dp e^{-\frac{p^2}{2}}$ is the complementary error function, and the equality holds by Part (1) of Proposition 4.4. The right side of (4.23) is finite by our conditions on g_λ , and since the denominator is approximately $\int_\Sigma ds h(s)$ for small λ , and thus bounded away from zero for $\lambda < 1$.

For $m = 2$, we write

$$\begin{aligned} \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\left(\sum_{n=0}^{\tilde{n}_1} g_\lambda(S_{\tau_n}) \right)^2 \right] &= \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} g_\lambda^2(S_{\tau_n}) + 2 \sum_{n < m}^{\tilde{n}_1} g_\lambda(S_{\tau_m}) g_\lambda(S_{\tau_n}) \right] \\ &= \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} g_\lambda^2(S_{\tau_n}) + 2 g_\lambda(S_{\tau_n}) \mathbb{E}^{(\lambda)} \left[\sum_{m=n+1}^{\tilde{n}_1} g_\lambda(S_{\tau_m}) \mid \tilde{\mathcal{F}}_{\tau_n^-} \right] \right] \\ &\leq \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} g_\lambda^2(S_{\tau_n}) + 2 g_\lambda(S_{\tau_n}) (U^{(\lambda)} g_\lambda)(S_{\tau_n}) \right]. \end{aligned} \quad (4.24)$$

The σ -algebra $\tilde{\mathcal{F}}_{\tau_n^-}$ contains the state S_{τ_n} since a.s. $\lim_{t \nearrow \tau_n} S_t = S_{\tau_n}$. However, by Part (3) of Proposition 4.3, the binary component Z_{τ_n} has probabilities $h(S_{\tau_n})$ and $1 - h(S_{\tau_n})$ for being 1 or 0 respectively given the information $\tilde{\mathcal{F}}_{\tau_n^-}$. By these observations and the strong Markov property at the time τ_n we get the first equality below

$$\mathbb{E}^{(\lambda)} \left[\sum_{m=n+1}^{\tilde{n}_1} g_\lambda(S_{\tau_m}) \mid \tilde{\mathcal{F}}_{\tau_n^-} \right] = \tilde{\mathbb{E}}_{\delta_{S_{\tau_n}}}^{(\lambda)} \left[\sum_{m=1}^{\tilde{n}_1} g_\lambda(S_{\tau_m}) \right] \leq (U^{(\lambda)} g_\lambda)(S_{\tau_n}), \quad (4.25)$$

where $\tilde{\delta}_s = \chi(z=0)(1-h(s))\delta_s + \chi(z=1)h(s)\delta_s$ is the splitting of the δ -distribution as $s \in \Sigma$. The inequality is the observation (4.21).

We can apply (4.23) with $f = g_\lambda^2 + 2g_\lambda U^{(\lambda)}g_\lambda$ to bound the right side of (4.24). Clearly, the contribution from g_λ^2 is not a problem, since $g_\lambda^2(x, p) \leq Cg_\lambda(x, p)$. For $g_\lambda U^{(\lambda)}(g_\lambda)$, there are constants such that

$$\begin{aligned} g_\lambda(p)(U^{(\lambda)}g_\lambda)(p) &\leq (\text{const}) \frac{1 + \log(1 + |p|)}{|p|^2}, & \text{for all } |p| \leq \lambda^{-1}, \lambda < 1 \\ g_\lambda(p)(U^{(\lambda)}g_\lambda)(p) &\leq (\text{const}) \frac{1 + \log(1 + |\lambda^{-1}|)}{|\lambda|^{-2}}, & \text{for all } |p| > \lambda^{-1}, \lambda < 1. \end{aligned}$$

We have applied Theorem 4.9 along with our conditions on g_λ to get

$$(U^{(\lambda)}g_\lambda)(x, p) \leq c + c \int_0^{|p|} dp' p' |g_\lambda(p')| \leq c + c' \int_0^{|p|} dp' \frac{p'}{1 + |p'|^2} \leq c + c'' \log(1 + |p|),$$

for some $c, c', c'' > 0$ and all $\lambda < 1$ and $|p| \leq \lambda^{-1}$. Similarly, we can obtain the analogous inequality for the domain $|p| > \lambda^{-1}$.

Now we sketch the proof for the general case $m > 2$. For $\epsilon_j \in \{<, =\}$, $j < m$ let the set $\ell^{(\mathbf{n}_1)}(\epsilon_1, \dots, \epsilon_{m-1})$ be the collection of all $(r_1, \dots, r_m) \in [0, \tilde{n}_1]^m$ with $r_1 \epsilon_1 r_2 \dots \epsilon_{m-1} r_m$. Also define,

$$f_{(\epsilon_1, \dots, \epsilon_{m-1})} = A_{\epsilon_1} \cdots A_{\epsilon_{m-1}} g_\lambda,$$

where A_+, A_- are linear operations in which A_+ is multiplication by g_λ and $A_- = A_+ U^{(\lambda)}$. We can write

$$\tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\left(\sum_{n=0}^{\tilde{n}_1} g_\lambda(\sigma_{r_M}) \right)^m \right] = \text{Lin. comb. over } (\epsilon_1 \cdots \epsilon_{m-1}) \text{ of } \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{\ell^{(\mathbf{n}_1)}(\epsilon_1, \dots, \epsilon_{m-1})} g_\lambda(\sigma_{r_1}) \cdots g_\lambda(\sigma_{r_m}) \right],$$

for $\tilde{\sigma}_n = \tilde{S}_{\tau_n}$. However,

$$\tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{\ell^{(\mathbf{n}_1)}(\epsilon_1, \dots, \epsilon_{m-1})} g_\lambda(\sigma_{r_1}) \cdots g_\lambda(\sigma_{r_m}) \right] \leq \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} f_{(\epsilon_1, \dots, \epsilon_{m-1})}(\sigma_n) \right], \quad (4.26)$$

since we can write the difference between the right and left side of (4.26) as a sum of positive terms $\mathbf{c}_{v-1} - \mathbf{c}_v$ indexed by $v \in [1, m-1]$, where

$$\mathbf{c}_v = \tilde{\mathbb{E}}_\nu^{(\lambda)} \left[\sum_{\ell^{(\mathbf{n}_1)}(\epsilon_1, \dots, \epsilon_v)} g_\lambda(\sigma_{r_1}) \cdots g_\lambda(\sigma_{r_v}) f_{(\epsilon_{v+1}, \dots, \epsilon_{m-1})}(\sigma_{r_{v+1}}) \right].$$

When ϵ_{v-1} is $=$, then \mathbf{c}_{v-1} and \mathbf{c}_v are identically equal. When ϵ_{v-1} is $<$, then the difference

$\mathbf{c}_v - \mathbf{c}_{v-1}$ is equal to

$$\begin{aligned}
& \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{\ell(\mathbf{n}_1)(\epsilon_1, \dots, \epsilon_{v-1})} g_{\lambda}(\sigma_{r_1}) \cdots g_{\lambda}(\sigma_{r_{v-1}}) \left(g_{\lambda}(\sigma_{r_v}) \sum_{n=r_v+1}^{\tilde{n}_1} f_{(\epsilon_{v+1}, \dots, \epsilon_{m-1})}(\sigma_{r_n}) - f_{(\epsilon_v, \dots, \epsilon_{m-1})}(\sigma_{r_v}) \right) \right] \\
&= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{\ell(\mathbf{n}_1)(\epsilon_1, \dots, \epsilon_{v-1})} g_{\lambda}(\sigma_{r_1}) \cdots g_{\lambda}(\sigma_{r_{v-1}}) \right. \\
&\quad \cdot \left(g_{\lambda}(\sigma_{r_v}) \tilde{\mathbb{E}}^{(\lambda)} \left[\sum_{n=r_v+1}^{\tilde{n}_1} f_{(\epsilon_{v+1}, \dots, \epsilon_{m-1})}(\sigma_{r_n}) \middle| \mathcal{F}_{S_{\tau_{r_v}^-}} \right] - f_{(\epsilon_v, \dots, \epsilon_{m-1})}(\sigma_{r_v}) \right) \left. \right] \\
&\leq \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{\ell(\mathbf{n}_1)(\epsilon_1, \dots, \epsilon_{v-1})} g_{\lambda}(\sigma_{r_1}) \cdots g_{\lambda}(\sigma_{r_{v-1}}) \left(g_{\lambda}(\sigma_{r_v}) (U^{(\lambda)} f_{(\epsilon_v, \dots, \epsilon_{m-1})})(\sigma_{r_v}) - f_{(\epsilon_v, \dots, \epsilon_{m-1})}(\sigma_{r_v}) \right) \right] \\
&= 0,
\end{aligned}$$

where the second relation is an inequality rather than an equality because $U^{(\lambda)} f_{(\epsilon_v, \dots, \epsilon_{m-1})}(\sigma_{r_v})$ over-counts the sum $\tilde{\mathbb{E}}^{(\lambda)} \left[\sum_{n=r_v+1}^{\tilde{n}_1} f_{(\epsilon_{v+1}, \dots, \epsilon_{m-1})}(\sigma_{r_n}) \middle| \mathcal{F}_{S_{\tau_{r_v}^-}} \right]$ in the event that $\zeta_{r_v} = 0$, analogously to (4.25).

Thus we are left to bound $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} f_{(\epsilon_1, \dots, \epsilon_{m-1})}(\sigma_n) \right]$. The worst case scenario is when all the ϵ_j are $<$, since mere multiplication by $g_{\lambda}(p)$ introduces more decay for large $|p|$. By our conditions on g_{λ} and $m-1$ applications of Theorem 4.9,

$$((U^{(\lambda)})^{m-1} g_{\lambda})(x, p) \leq c^{m-1} \frac{(1 + \log(1 + |p|))^{m-1}}{1 + |p|^2}, \quad |p| \leq \lambda^{-1},$$

and we get another bound for $|p| > \lambda^{-1}$ which is smaller than a fixed multiple of λ^{-2m-1} for all $\lambda < 1$. Applying the inequality (4.23) then we attain the bound. \square

The proof of the following proposition is analogous (although simpler) to that of Proposition 4.14 in the next section, which concerns the case of the oscillatory (non-integrable) function $g(x, p) = \frac{dV}{dx}(x)$.

Proposition 4.11. *Let $g : \Sigma \rightarrow \mathbb{R}$ satisfy that $|g(x, p)| \leq \frac{C}{1+|p|^2}$ for some $C > 0$ and all $(x, p) \in \Sigma$.*

1. *For any $m \in \mathbb{N}$, there is a $C > 0$ such that*

$$\sup_{\lambda \leq 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t \leq R_1} \left(\int_0^t dr g(S_r) \right)^{2m} \right] < C.$$

2. *There is a $C > 0$ such that*

$$\sup_{\lambda \leq 1} \tilde{\mathbb{E}}_{(x, p, z)}^{(\lambda)} \left[\left| \int_0^{R_1} dr g(S_r) \right| \right] < C(1 + \log(1 + |p|)).$$

4.4 Inequalities for the momentum drift

This section is focused on various inequalities relating to the expectation and higher moments of $\int_{\varsigma_1}^{\varsigma_2} dr \frac{dV}{dx}(X_r)$ for certain hitting times $\varsigma_1 \leq \varsigma_2$. One of the main points is Part (1) of Proposition 4.14, which states that

$$\sup_{\lambda \leq 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t \leq R_1} \left| \int_0^t dr \frac{dV}{dx}(X_r) \right|^m \right] < \infty$$

for each $m > 0$. This implies that there will typically not be large fluctuations in the momentum drift $D_t = \int_0^t dr \frac{dV}{dx}(X_r)$ for t between successive returns R_n, R_{n+1} to the atom (i.e. during a single life cycle).

The first two parts in the lemma below follow from the conservation of energy and the quadratic formula and do not depend on the potential being periodic. The third part of Lemma 4.12 is a statement about mixing on the torus. If the particle begins with a high momentum $|P_0| \gg 1$ and is stopped at a random exponential time τ , then the distribution on the torus $\mathbb{T} = [0, 1)$ at the stopping time will be roughly uniform—even in the presence of the bounded periodic potential $V(x)$

Lemma 4.12. *Let (X_t, P_t) evolve according to the Hamiltonian $H(x, p) = \frac{1}{2}p^2 + V(x)$, for a positive potential $V(x)$ with $\sup_x V(x) < \infty$ and $\sup_x \left| \frac{dV}{dx}(x) \right| < \infty$. If the initial momentum has $|P_0|^2 > 4\bar{V}$, then the difference $P_t - P_0 = \int_0^t dr \frac{dV}{dx}(X_r)$ satisfies*

1. $\sup_{t \in \mathbb{R}^+} \left| \int_0^t dr \frac{dV}{dx}(X_r) \right| \leq 2 \sup_x V(x) |P_0|^{-1}$, and
2. $\left| \int_0^t dr \frac{dV}{dx}(X_s) - \frac{V(X_t) - V(X_0)}{P_0} \right| \leq 2t \sup_x \left| \frac{dV}{dx}(x) \right| \sup_x V(x) |P_0|^{-2}$.
3. Suppose further that $V(x)$ has period one. If τ is exponentially distributed with mean \mathbf{r}^{-1} and $F : \mathbb{T} \rightarrow \mathbb{R}$ is a function of the torus and bounded, then

$$\left| \mathbb{E}_{(X_0, P_0)} [F(X_\tau)] - \int_{\mathbb{T}} dx F(x) \right| \leq \mathbf{r} \|F\|_{\infty} |P_0|^{-1} + O(|P_0|^{-2}).$$

Proof.

Part (1): Since $|P_0|^2 > 4 \sup_x V(x)$, the momentum P_t will not change sign at any time. By the conservation of energy

$$\frac{1}{2} |P_0 + (P_t - P_0)|^2 - \frac{1}{2} P_0^2 = -V(X_t) + V(X_0).$$

Using the quadratic formula and that P_t, P_0 have the same sign,

$$|P_t - P_0| = \left| |P_0| - (P_0^2 + 2V(X_0) - 2V(X_t))^{\frac{1}{2}} \right| \leq \left| \frac{1}{2} \int_0^{2V(X_0) - 2V(X_t)} dw (P_0^2 + w)^{-\frac{1}{2}} \right| < \frac{2 \sup_x V(x)}{|P_0|},$$

since $(P_0^2 + w)^{-\frac{1}{2}} \leq \sqrt{2} |P_0|^{-1} < 2 |P_0|^{-1}$ for $|w| \leq \frac{1}{2} P_0^2$.

Part (2): With the identity $V(X_t) - V(X_0) = \int_0^t dr \frac{dV}{dx}(X_r) P_r$, then

$$\begin{aligned} \left| \int_0^t dr \frac{dV}{dx}(X_s) - \frac{V(X_t) - V(X_0)}{P_0} \right| &\leq \int_0^t dr \left| \frac{dV}{dx}(X_r) \left(1 - \frac{P_r}{P_0}\right) \right| \\ &\leq t |P_0|^{-1} \sup_x \left| \frac{dV}{dx}(x) \right| \sup_r |P_r - P_0| \leq 2t \sup_x \left| \frac{dV}{dx}(x) \right| \sup_x V(x) |P_0|^{-2}, \end{aligned}$$

where we applied Part (1) for the last inequality.

Part (3): Let $d_s : \mathbb{T} \rightarrow \mathbb{R}^+$ be the density of the particle at time τ on the torus starting from the point $s = (X_0, P_0) \in \Sigma$. Then,

$$\mathbb{E}_s[F(X_\tau)] = \int_{\mathbb{T}} dx d_s(x) F(x). \quad (4.27)$$

This leads to the simple bound

$$|\mathbb{E}_s[F(X_\tau)] - \int_{\mathbb{T}} dx F(x)| \leq \|F\|_\infty \|d_s - 1\|_1.$$

Thus it is sufficient to show to bound the 1-norm of $d_s - 1$, and, in fact, our bounds can be made in the supremum norm.

Notice that d_s can be written as

$$d_s(a) = \sum_{n=1}^{\infty} |P_{t_n(a)}|^{-1} \mathbf{r} e^{-\mathbf{r} t_n(a)},$$

where $t = t_1(a), t_2(a), \dots$ are the periodic sequence of times for which $X_t \bmod(1) = a$. These times will exist for every $a \in \mathbb{T}$ as long as $H(X_0, P_0) > \sup_x V(x)$.

If $4 \sup_x V(x) \leq P_0^2$, then $|P_t - P_0| \leq 2(\sup_x V(x)) |P_0|^{-1}$ by Part (1). Thus for large momentum $|P_0| \gg (\sup_x V(x))^{\frac{1}{2}}$, P_r is nearly constant and the hitting times $t_n(a)$ will be close to the sequence of times $t = t'_n(a)$ at which $X_0 + t P_0 \bmod(1) = a$ for a time period at least on the order of $t^{\frac{1}{2}}$. The period Δ such that $t'_n(a) - t'_{n-1}(a) = \Delta$ should thus be close to $\frac{1}{|P_0|}$. When $|P_0|$ is large enough so that $|P_t - P| \leq 2 \sup_x V(x) |P_0|^{-1} < \frac{1}{2} |P_0|$, then clearly $\Delta \leq \frac{2}{|P_0|}$, and

$$\begin{aligned} \left| \Delta - \frac{1}{|P_0|} \right| &\leq \frac{1}{|P_0|} \left| \int_0^\Delta dr P_0 - \int_0^{\frac{1}{|P_0|}} dr P_0 \right| \leq \frac{1}{|P_0|} \left(\int_0^\Delta dr |P_r - P_0| + \int_0^{\frac{1}{|P_0|}} dr |P_r - P_0| \right) \\ &< \frac{6 \sup_x V(x)}{|P_0|^3}. \end{aligned} \quad (4.28)$$

The difference between the first crossing-times $|t'_1(a) - t_1(a)|$ of the point a can be similarly bounded.

Using the triangle inequality

$$\begin{aligned} |d_s(a) - 1| &\leq \left| d_s(a) - \frac{1}{|P_0|} \sum_{n=1}^{\infty} \mathbf{r} e^{-\mathbf{r} t'_n(a)} \right| + \left| \frac{1}{|P_0|} \sum_{n=1}^{\infty} \mathbf{r} e^{-\mathbf{r} t'_n(a)} - \frac{1}{|P_0|} \sum_{n=1}^{\infty} \mathbf{r} e^{-\mathbf{r} t_n(a)} \right| \\ &\quad + \left| \frac{1}{|P_0|} \sum_{n=1}^{\infty} \mathbf{r} e^{-\mathbf{r} t_n(a)} - 1 \right| \\ &\leq \frac{2\mathbf{r}}{|P_0|} + O\left(\frac{1}{|P_0|^2}\right), \end{aligned} \quad (4.29)$$

where the last inequality follows by further computation using the inequalities above. For instance, we can bound the first term on the right as

$$|d_s(a) - \frac{1}{|P_0|} \sum_{n=1}^{\infty} \mathbf{r} e^{-\mathbf{r} t'_n(a)}| \leq \sup_n \left| 1 - \frac{P_0}{P_{t'_n(a)}} \right| e^{-\mathbf{r} t'_1(a)} \frac{1}{|P_0| \Delta} \frac{\mathbf{r} \Delta}{1 - e^{-\mathbf{r} \Delta}} \leq \frac{2 \sup_x V(x)}{P_0^2}.$$

□

Define the functions $\mathbf{C}_n^{(\lambda)} : \Sigma \rightarrow \mathbb{R}$,

$$\begin{aligned} \mathbf{C}_0^{(\lambda)}(s) &= \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\chi(Z_{\tau_1} = 0) \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right], \\ \mathbf{C}_n^{(\lambda)}(s) &= \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\left(\chi(Z_{\tau_1} = 0) \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(s) \right)^{2n} \right], \end{aligned}$$

where τ_1, τ_2 are the first two partition times, $n \in \mathbb{N}$, and $\tilde{\delta}_s = (1 - h(s))\delta_{(s,0)} + h(s)\delta_{(s,1)}$ (i.e. the splitting of the δ -distribution at s). The presence of the factor $\chi(Z_{\tau_1} = 0)$ in the above definitions is a small technical precaution, and if $\chi(Z_{\tau_1} = 0)$ is removed in the formula for $\mathbf{C}_0^{(\lambda)}(s)$, then we have

$$\tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] = \mathbb{E}_s^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] = \mathbb{E}_s^{(\lambda)} \left[\int_0^\infty dt t e^{-t} \frac{dV}{dx}(S_t) \right].$$

Lemma 4.13. *For any $n > 1$, there exists a $C > 0$ such that for all $\lambda < 1$, then*

1. $|\mathbf{C}_n^{(\lambda)}(x, p)| \leq C \left(\frac{1}{1+|p|^{2n}} \vee \lambda^{2n} \right)$, and
2. $|\mathbf{C}_0^{(\lambda)}(x, p)| \leq C \left(\frac{1}{1+|p|^2} \vee \lambda \right)$.

Proof.

Part (1): For $v = 2m$, $\mathbf{C}_m^{(\lambda)}(s)$ is smaller than

$$\mathbf{C}_m^{(\lambda)}(s) \leq \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^v \right] = \mathbb{E}_s^{(\lambda)} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^v \right],$$

where the equality is since the initial distribution $\tilde{\delta}_s$ is the splitting of δ_s , and the argument of the expectations only depends on the pre-split statistics.

We will bound $\mathbb{E}_s^{(\lambda)} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^v \right]$ for $s = (x, p)$ in the three regimes of

- (i). arbitrary p ,
- (ii). $1 \ll |p| \leq \lambda^{-1}$, and
- (iii). $\lambda^{-1} < |p|$.

In the third regime, the absolute value of the momentum is generally drifting downwards and there is also a higher frequency of collisions.

(i). For arbitrary $s \in \Sigma$, we have

$$\mathbb{E}_s^{(\lambda)} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^v \right] \leq \sup_{x \in \mathbb{T}} \left| \frac{dV}{dx}(x) \right|^v \mathbb{E}[\tau_2 - \tau_1] \leq v! \sup_{x \in \mathbb{T}} \left| \frac{dV}{dx}(x) \right|^v, \quad (4.30)$$

since $\tau_2 - \tau_1$ is a mean-1 exponential.

(ii). Next, we consider $s = (x, p)$ for the regime $1 \ll |p| < \lambda^{-1}$.

$$\begin{aligned} \mathbb{E}_s^{(\lambda)} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^v \right] &\leq \sup_{\substack{s'=(x,p) \\ |p| \in [\frac{3}{4}|p|, \frac{5}{4}|p|]}} \mathbb{E}_{s'}^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right] \\ &\quad + \left(v! \sup_{x \in \mathbb{T}} \left| \frac{dV}{dx}(x) \right|^v \right)^{\frac{1}{2}} \mathbb{P}_s^{(\lambda)} \left[|P_r| \notin \left[\frac{3}{4}|p|, \frac{5}{4}|p| \right], \text{ for some } r \leq \tau_1 \right]^{\frac{1}{2}}. \end{aligned} \quad (4.31)$$

For the second inequality, we have split the expectation into two terms corresponding to whether the event $|P_r| \in [\frac{3}{4}|p|, \frac{5}{4}|p|]$ for all $r \in [0, \tau_1]$ occurred or not. In the case that the event occurs, then we get the first term on the right side in which we have used the Markov property and relabeled τ_2 and as the first partition time τ_1 (starting from zero rather than τ_1). For the complementary event, we applied Cauchy-Schwarz and the inequality (4.30).

To bound the second term on the right, we may apply Chebyshev's inequality,

$$\begin{aligned} \mathbb{P}_s^{(\lambda)} \left[|P_r| \notin \left[\frac{3}{4}|p|, \frac{5}{4}|p| \right], \text{ for some } r \leq \tau_1 \right] &\leq \left(\frac{4}{|p|} \right)^w \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \rho \wedge \tau_1} |P_r - p|^w \right] \\ &\leq \left(\frac{8}{|p|} \right)^w \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq r \leq \rho \wedge \tau_1} |J_r|^w \right] + w! \left(\frac{8}{|p|} \right)^w \sup_{x \in \mathbb{T}} \left| \frac{dV}{dx}(x) \right|^w, \end{aligned} \quad (4.32)$$

where $w \geq 1$, ρ is the first jump time $|P_r|$ leaves $[\frac{3}{4}|p|, \frac{5}{4}|p|]$, and $J_r = P_r - p - \int_0^t dr \frac{dV}{dx}(X_r)$ is the sum of the momentum jumps up to time r . For the second inequality, we have used $(x + y)^w \leq 2^w(x^w + y^w)$ and (4.30) to bound the contribution of the potential drift. The absolute values of the jumps $|J_{t_n} - J_{t_{n-1}}|$ have uniformly controlled Gaussian tails and occur with Poisson rate $\mathcal{E}_\lambda(P_r) \leq \mathcal{E}_\lambda(2\lambda^{-1})$ for $r \leq \rho$, and thus the expectation of $\sup_{0 \leq r \leq \rho \wedge \tau_1} |P_r - p|^{2v}$ is finite. It follows that (4.32) decays super-polynomially fast for $|p| \gg 1$.

Now we bound the first term on the right side of (4.31). Let ς be the hitting time that S_t jumps out of the set $A_p = \{(x', p') \in \Sigma \mid ||p'| - |p|| \leq \frac{1}{2}|p|\}$ starting from some $s' = (x, p)$ for $|p| \in [\frac{3}{4}|p|, \frac{5}{4}|p|]$. Define the times $t'_n = t_n \wedge \tau_1 \wedge \varsigma$, where t_n is the time of the n th momentum jump. By writing,

$$\int_0^{\tau_1} dr \frac{dV}{dx}(X_r) = \sum_{n=0}^{\infty} \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) + \chi(\varsigma \leq \tau_1) \int_{\varsigma}^{\tau_1} dr \frac{dV}{dx}(X_r),$$

we can apply the triangle inequality to get

$$\begin{aligned} \mathbb{E}_{s'}^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right]^{\frac{1}{v}} &\leq \mathbb{E}_{s'}^{(\lambda)} \left[\left(\sum_{n=0}^{\infty} \left| \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) \right| \right)^v \right]^{\frac{1}{v}} + \mathbb{E}_{s'}^{(\lambda)} \left[\chi(\varsigma \leq \tau_1) \left| \int_{\varsigma}^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right]^{\frac{1}{v}} \\ &\leq \mathbb{E}_{s'}^{(\lambda)} \left[\left(\sum_{n=0}^{\infty} \left| \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) \right| \right)^v \right]^{\frac{1}{v}} + \left((2v)! \sup_{x \in \mathbb{T}} \left| \frac{dV}{dx}(x) \right|^{2v} \right)^{\frac{1}{2v}} \mathbb{P}_{s'}^{(\lambda)} [\varsigma \leq \tau_1]^{\frac{1}{2v}}, \end{aligned} \quad (4.33)$$

where the second inequality follows by Cauchy-Schwarz and that τ_1 is a mean-1 exponential. The probability $\mathbb{P}_{s'}^{(\lambda)}[\varsigma \leq \tau_1]$ decays faster than any polynomial by the same reasoning as for (4.32). The first term on the right side of (4.33) is bounded by

$$\mathbb{E}_{s'}^{(\lambda)} \left[\left(\sum_{n=0}^{\infty} \left| \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) \right| \right)^v \right] \leq \left(\frac{4 \sup_{x \in \mathbb{T}} V(x)}{|p|} \right)^v \mathbb{E}_{s'}^{(\lambda)} [\mathcal{N}_{\varsigma}^v], \quad (4.34)$$

where \mathcal{N}_t is the number of collisions up to time t . The above inequality uses the definition of the t'_n 's to conclude that for each n , either $t'_n = t'_{n+1}$ so that $\int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) = 0$, or $|P_{t'_n}| \geq \frac{1}{2}|p|$ so that we can apply Part (1) of Lemma 4.12 to bound $|\int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r)|$. The counting process \mathcal{N}_t has Poisson rate $\mathcal{E}_{\lambda}(P_t)$ at time t . Up to time ς , we have that $\mathcal{E}_{\lambda}(P_t) \leq \mathcal{E}_{\lambda}(2\lambda^{-1}) \leq \mathbf{r}$ for a constant \mathbf{r} which is guaranteed to be finite by Part (1) of Proposition 3.1.

$$\mathbb{E}_{s'}^{(\lambda)} [\mathcal{N}_{\varsigma}^v] \leq \mathbb{E}[(N'_{\tau})^v] = \frac{1}{1+\mathbf{r}} \sum_{n=0}^{\infty} n^v \left(\frac{\mathbf{r}}{1+\mathbf{r}} \right)^n,$$

where N'_t is a Poisson process with rate \mathbf{r} and τ is mean-1, exponentially distributed, and independent of N'_t . The first inequality can be seen by a construction $N'_{\tau} \approx \mathcal{N}_{\varsigma} + \mathcal{N}'_{\tau}$ for a jump process \mathcal{N}'_{τ} with Poisson jump rate $\mathbf{r} - \mathcal{E}_{\lambda}(P_t)$ for $t \leq \varsigma$ and rate \mathbf{r} for $t > \varsigma$ whose jumps are decided independently of the jumps for \mathcal{N}_{τ} .

(iii). Let $\vartheta = \tau_1 \wedge \vartheta'$ where ϑ' is the hitting time that the absolute value of the momentum $|P_t|$ jumps below λ^{-1} . The hitting time ϑ' is finite, and, in fact, has an expectation that is bounded by a multiple of λ^{-1} independently of the initial momentum $|p| > \lambda^{-1}$. However, the details for these points do not matter for this proof. Let φ_s be the distribution on $\mathbb{T} \times [-\lambda^{-1}, \lambda^{-1}]$ for $(X_{\vartheta'}, P_{\vartheta'})$ starting from $s \in \Sigma$. By the triangle inequality and the strong Markov property

$$\mathbb{E}_s^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right]^{\frac{1}{v}} \leq \mathbb{E}_s^{(\lambda)} \left[\left| \int_0^{\vartheta} dr \frac{dV}{dx}(X_r) \right|^v \right]^{\frac{1}{v}} + \mathbb{E}_{\varphi_s}^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right]^{\frac{1}{v}}. \quad (4.35)$$

For the first term on the right side on (4.35), we can write

$$\begin{aligned} \mathbb{E}_s^{(\lambda)} \left[\left| \int_0^{\vartheta} dr \frac{dV}{dx}(X_r) \right|^v \right] &= \mathbb{E}_s^{(\lambda)} \left[\left| \sum_{n=1}^{\mathcal{N}_{\vartheta}} \int_{t_{n-1}}^{t_n} dr \frac{dV}{dx}(X_r) \right|^v \right] \\ &\leq 2^v \left(\sup_x V(x) \right)^v \mathbb{E}_s^{(\lambda)} \left[\left| \sum_{n=1}^{\mathcal{N}_{\vartheta}} \frac{1}{|P_{t_n}^-|} \right|^v \right] \leq C \lambda^v \mathbb{E}_s^{(\lambda)} \left[\left| \sum_{n=1}^{\mathcal{N}_{\vartheta}} \frac{1}{\mathcal{E}_{\lambda}(P_{t_n}^-)} \right|^v \right]. \end{aligned} \quad (4.36)$$

The first inequality is Part (1) of Lemma 4.12, which is applied for the Hamiltonian evolution on each interval $[t_{n-1}, t_n]$. The C for the second inequality is guaranteed by Part (2) of Proposition 3.1. Notice that $\sum_{n=1}^{\mathcal{N}_r} \frac{1}{\mathcal{E}_{\lambda}(P_{t_n}^-)} - r$ is a martingale with predictable quadratic variation $\int_0^r ds \frac{1}{\mathcal{E}_{\lambda}(P_s)}$, since the counting process \mathcal{N}_r has jump rate $\mathcal{E}_{\lambda}(P_r)$. Since $\vartheta \leq \tau_1$ and by the triangle inequality,

$$\begin{aligned} \mathbb{E}_s^{(\lambda)} \left[\left| \sum_{n=1}^{\mathcal{N}_{\vartheta}} \frac{1}{\mathcal{E}_{\lambda}(P_{t_n}^-)} \right|^v \right]^{\frac{1}{v}} &\leq \mathbb{E}_s^{(\lambda)} \left[\left| \sum_{n=1}^{\mathcal{N}_{\tau_1}} \frac{1}{\mathcal{E}_{\lambda}(P_{t_n}^-)} - \tau_1 \right|^v \right]^{\frac{1}{v}} + \mathbb{E}_s^{(\lambda)} [\tau_1^v]^{\frac{1}{v}} \\ &\leq C' \mathbb{E}_s^{(\lambda)} \left[\left| \int_0^{\tau_1} ds \frac{1}{\mathcal{E}_{\lambda}(P_s)} \right|^v \right]^{\frac{1}{v}} + \mathbb{E}_s^{(\lambda)} \left[\sup_{1 \leq n \leq \mathcal{N}_{\tau_1}} \frac{1}{\mathcal{E}_{\lambda}(P_{t_n}^-)} \right]^{\frac{1}{v}} + \mathbb{E}_s^{(\lambda)} [\tau_1^v]^{\frac{1}{v}} \\ &\leq (C' + 1)(1 + \lambda)8(v!)^{\frac{1}{v}} + (v!)^{\frac{1}{v}}, \end{aligned}$$

where the constant C' comes from Rosenthal's inequality (see e.g. [11], Lemma 2.1), and third inequality is because $\mathcal{E}_\lambda(p) \geq \frac{1}{8(1+\lambda)}$ by Part (1) of Proposition 3.1 and since τ_1 is exponential with mean one.

For the second term on the right side of (4.35), we can apply our results (i) and (ii) above to guarantee the existence of a $C > 0$ such that

$$\begin{aligned} \mathbb{E}_{\varphi_s}^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right|^v \right] &\leq C \int_{\Sigma} d\varphi_s(x', p') \frac{1}{1 + |p'|^v} \\ &\leq C \sup_{|p'| > \frac{1}{\lambda}} \frac{\int_{[-\frac{1}{\lambda}, \frac{1}{\lambda}]} dp'' \frac{1}{1 + |p''|^v} \mathcal{J}_\lambda(p', p'')}{\int_{[-\frac{1}{\lambda}, \frac{1}{\lambda}]} dp'' \mathcal{J}_\lambda(p', p'')} = O(\lambda^v) + C \sup_{|p'| > \frac{1}{\lambda}} \frac{\int_{[-\frac{1}{2\lambda}, \frac{1}{2\lambda}]} dp'' \frac{1}{1 + |p''|^v} \mathcal{J}_\lambda(p', p'')}{\int_{[-\frac{1}{\lambda}, \frac{1}{\lambda}]} dp'' \mathcal{J}_\lambda(p', p'')}, \end{aligned} \quad (4.37)$$

where the third expression should be understood as the supremum over all $|p'| > \lambda^{-1}$ for the expectation for $\frac{1}{1 + |P_\tau|^v}$ conditioned on $p' = P_{\tau-}$. The final term in (4.37) is super-polynomially small by Part (7) of Proposition 3.1 and Chebyshev's inequality.

Part (2): By (4.30), we have an upper bound for $\sup_{(x,p)} |\mathbf{C}_0^{(\lambda)}(x, p)|$. For p in the regime $1 \ll |p| \leq \lambda^{-1}$, define ς as in Part (1). Also define t_n as the sequence of collision times starting after τ_1 with $t_0 = \tau_1$, and $t'_n = t_n \wedge \varsigma \wedge \tau_2$. The difference between $\mathbf{C}_0^{(\lambda)}(s)$ and $\mathbb{E}_s^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right]$ is small when $|p| \gg 1$, since

$$\begin{aligned} \left| \mathbb{E}_s^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] - \mathbf{C}_0^{(\lambda)}(s) \right| &= \left| \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] - \mathbf{C}_0^{(\lambda)}(s) \right| \\ &= \left| \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[\chi(z_{\tau_1} = 1) \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] \right| \leq \tilde{\mathbb{E}}_{\tilde{\delta}_s} \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^2 \right]^{\frac{1}{2}} \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)} \left[h^2(S_{\tau_1}) \right]^{\frac{1}{2}} \\ &= \mathbb{E}_s \left[\left| \int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right|^2 \right]^{\frac{1}{2}} \mathbb{E}_s^{(\lambda)} \left[h^2(S_{\tau_1}) \right]^{\frac{1}{2}} \leq 2^{\frac{1}{2}} \sup_x \left| \frac{dV}{dx}(x) \right|^{\frac{1}{2}} \mathbb{E}_s^{(\lambda)} \left[h^2(S_{\tau_1}) \right]^{\frac{1}{2}}, \end{aligned}$$

The first and the third equalities have used that $\mathbb{E}_s^{(\lambda)} = \tilde{\mathbb{E}}_{\tilde{\delta}_s}^{(\lambda)}$. The first inequality is Cauchy-Schwarz, and the second inequality uses that $\tau_2 - \tau_1$ is a mean-1 exponential. The function $h(s) \leq 1$ has compact support, and there is a $c > 0$ such that $\mathbb{E}_{(x,p)}^{(\lambda)} [h(S_{\tau_1})] \leq ce^{-\lambda^{-1}} \vee e^{-|p|}$. In fact, the bound can be given a Gaussian form by the Gaussian tails for the jump rates (1.2).

By the above remarks, we may work with $\mathbb{E}_s^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right]$, since the error is $O(|p|^{-2})$. Now we will show that the difference of this term with the expression $\frac{1}{p} \mathbb{E}_s^{(\lambda)} [V(X_{\tau_2}) - V(X_{\tau_1})]$ is also $O(|p|^{-2})$. Similarly to Part (1), we can write

$$\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) = \sum_{n=0}^{\infty} \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) + \chi(\varsigma \leq \tau_2) \int_{\varsigma \vee \tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r).$$

The difference is bounded by

$$\begin{aligned} \left| \mathbb{E}_s^{(\lambda)} \left[\int_{\tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r) \right] - \frac{1}{p} \mathbb{E}_s^{(\lambda)} [V(X_{\tau_2}) - V(X_{\tau_1})] \right| &\leq \left(\frac{\sup_x V(x)}{|p|} + \sup_x \left| \frac{dV}{dx}(x) \right| \right) \mathbb{P}_s[\varsigma \leq \tau_2] \\ &\quad + \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} \left| \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) - \frac{V(X_{t'_{n+1}}) - V(X_{t'_n})}{P_{t'_n}} \right| \right] \\ &\quad + \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} |V(X_{t'_{n+1}}) - V(X_{t'_n})| \left| \frac{1}{P_{t'_n}} - \frac{1}{p} \right| \right], \end{aligned} \quad (4.38)$$

where \mathcal{N}_r , $r \geq \tau_1$ is the number of collision times t_n in the interval $(\tau_1, r]$, and the first term on the right side occurs by bounding the expectations of $\frac{V(X_{\tau_2}) - V(X_\varsigma)}{p}$ and $\chi(\varsigma \leq \tau_2) \int_{\varsigma \vee \tau_1}^{\tau_2} dr \frac{dV}{dx}(X_r)$. The above follows by adding and subtracting terms $\frac{V(X_{t'_{n+1}}) - V(X_{t'_n})}{P_{t'_n}}$ for $n \in [0, \mathcal{N}_\varsigma)$ and applying the triangle inequality. By the analysis in Part (1), $\mathbb{P}_s[\varsigma \leq \tau_2]$ decays super-polynomially in p .

$$\begin{aligned} & \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} \left| \int_{t'_n}^{t'_{n+1}} dr \frac{dV}{dx}(X_r) - \frac{V(X_{t'_{n+1}}) - V(X_{t'_n})}{P_{t'_n}} \right| \right] \\ & \leq 2 \sup_x \left| \frac{dV}{dx}(x) \right| \sup_x V(x) \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} (t'_{n+1} - t'_n) |P_{\tau_n}|^{-2} \right] \leq \frac{8}{|p|^2} \sup_x \left| \frac{dV}{dx}(x) \right| \sup_x V(x), \end{aligned}$$

where the second inequality uses that $|P_{\tau'_n}| \geq \frac{1}{2}|p|$, by definition, for $n \leq \mathcal{N}_\varsigma$, and also uses

$$\sum_{n=0}^{\mathcal{N}_\varsigma} t'_{n+1} - t'_n = \varsigma \leq \tau_2 - \tau_1 \quad \text{so that} \quad \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma} t'_{n+1} - t'_n \right] = \mathbb{E}_s^{(\lambda)}[\varsigma] \leq \mathbb{E}_s^{(\lambda)}[\tau_2 - \tau_1] = 1.$$

For the third term on the right side of (4.38),

$$\begin{aligned} & \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} |V(X_{t'_{n+1}}) - V(X_{t'_n})| \left| \frac{1}{P_{t'_n}} - \frac{1}{p} \right| \right] \\ & \leq \sup_x V(x) \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} \frac{|p - P_{t'_n}|}{|p P_{t'_n}|} \right] \leq \frac{4}{|p|^2} \left(\sup_x \left| \frac{dV}{dx}(x) \right| \right) \left(\sup_x V(x) \right) \mathbb{E}_s^{(\lambda)} \left[\sum_{n=0}^{\mathcal{N}_\varsigma-1} |p - P_{t'_n}| \right] \\ & \leq \frac{4}{|p|^2} \left(\sup_x \left| \frac{dV}{dx}(x) \right| \right) \left(\sup_x V(x) \right) \mathbb{E}_s^{(\lambda)} \left[\sup_{0 \leq r \leq \varsigma} |P_r - p|^2 \right]^{\frac{1}{2}} \mathbb{E}_s^{(\lambda)}[\mathcal{N}_\varsigma^2]^{\frac{1}{2}} = O(|p|^{-2}). \end{aligned}$$

The first inequality uses the definition of ς to conclude that $\frac{1}{2}|p| \leq |P_{t'_n}| \leq \frac{3}{2}|p|$ for $n \leq \mathcal{N}_\varsigma$, and the third is Cauchy-Schwarz. Arbitrary moments of \mathcal{N}_ς are finite by (4.34).

For the term $\frac{1}{p} \mathbb{E}_s^{(\lambda)}[V(X_{\tau_2}) - V(X_{\tau_1})]$, we just need to show $|\mathbb{E}_s^{(\lambda)}[V(X_{\tau_2}) - V(X_{\tau_1})]| = O(|p|^{-1})$. By the triangle inequality,

$$|\mathbb{E}_s^{(\lambda)}[V(X_{\tau_2}) - V(X_{\tau_1})]| \leq \left| \mathbb{E}_s^{(\lambda)}[V(X_{\tau_2})] - \int_{\mathbb{T}} dx V(x) \right| + \left| \mathbb{E}_s^{(\lambda)}[V(X_{\tau_1})] - \int_{\mathbb{T}} dx V(x) \right|.$$

The terms on the right side are similar, so we will study the second. Let us offer a reconstruction of the counting process \mathcal{N}_t for the number of collisions up to time ς as follows. Let N' be a Poisson clock with rate $\mathbf{r} = \mathcal{E}_\lambda(2\lambda^{-1})$ as in Part (1). The Poisson rate of jumps $\mathcal{E}_\lambda(P_t)$ for the process \mathcal{N}_t satisfies $\mathcal{E}_\lambda(P_t) \leq \mathbf{r}$ for times $t \leq \varsigma$. At each jump time $r_n \leq \varsigma$ for the Poisson process N' , we then flip an independent coin with weight $\mathbf{r}^{-1} \mathcal{E}_\lambda(P_{r_n})$ to determine if a jump for \mathcal{N}_t (i.e. a collision) occurred at time r_n . This recovers the statistics for \mathcal{N}_t . We then define $r'_n = r_n \wedge \tau_1$ for $n \leq N'_{\varsigma \wedge \tau_1}$. Conditioned on the past $\mathcal{F}_{r'_n}$ and the event $\tau_1 > r'_n$, the increment $r'_{n+1} - r'_n$ is exponentially distributed with mean $(1 + \mathbf{r})^{-1}$. When conditioned on the event $\tau_1 = r'_{n+1}$, the increment is exponential with mean 1.

We can rewrite the expectation $\mathbb{E}_s^{(\lambda)}[V(X_{\tau_1})]$ in terms of the event $\tau_1 = r'_n$ for $n \geq 1$ or $\tau_1 > \max_n r'_n$:

$$\begin{aligned} \mathbb{E}_s^{(\lambda)}[V(X_{\tau_1})] &= \mathbb{E}_s^{(\lambda)}[V(X_{\tau_1})\chi(\tau_1 > \max_n r'_n)] \\ &\quad + \mathbb{E}_s^{(\lambda)}\left[\sum_{n=0}^{\infty} \chi(\tau_1 = r'_{n+1}) \mathbb{E}_s^{(\lambda)}[V(X_{\tau_1}) \mid \mathcal{F}_{r'_n}, \tau_1 = r'_{n+1}]\right]. \end{aligned} \quad (4.39)$$

The first term on the right is smaller than $\sup_x V(x)$ times the probability of the event $\max_n r'_n \neq \tau_1$, which can also be phrased as the event that $\varsigma < \tau$. By the analysis in Part (1), $\mathbb{P}_s^{(\lambda)}[\varsigma < \tau]$ is super-polynomially small in $|p|$. Thus $\sum_{n=0}^{\infty} \mathbb{P}_s^{(\lambda)}[\tau_1 = r'_{n+1}]$ is super-polynomially close to 1. Since $r'_{n+1} - r'_n$ is exponentially distributed, by Part (3) of Lemma 4.12

$$\left| \mathbb{E}_s^{(\lambda)}[V(X_{\tau_1}) \mid \mathcal{F}_{r'_n}, \tau_1 = r'_{n+1}] - \int_{\mathbb{T}} dx V(x) \right| \leq (\text{const})|P_{r'_n}|^{-1} \leq 2(\text{const})|p|^{-1},$$

where the second inequality follows since $|P_{r'_n}| \geq \frac{1}{2}|p|$ by the definition of the r'_n , which are less than ς .

The analysis of the region $|p| \geq \lambda^{-1}$ follows from that of Part (1), and so the proof is complete. \square

Proposition 4.14.

1. For any $m \in \mathbb{N}$, there is a $C > 0$ such that

$$\sup_{\lambda \leq 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t \leq R_1} \left(\int_0^t dr \frac{dV}{dx}(X_r) \right)^{2m} \right] < C.$$

2. There is a $C > 0$ such that

$$\sup_{\lambda \leq 1} \tilde{\mathbb{E}}_{(x,p,z)}^{(\lambda)} \left[\left| \int_0^{R_1} dr \frac{dV}{dx}(X_r) \right| \right] < C(1 + \log(1 + |p|)).$$

3. As $\lambda \rightarrow 0$,

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \int_r^{R_2} dr' \frac{dV}{dx}(X_{r'}) \right] \\ = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(0)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \int_r^{R_2} dr' \frac{dV}{dx}(X_{r'}) \right] + O(\lambda^{\frac{1}{40}}). \end{aligned}$$

Proof.

Part (1): For $0 \leq t < R_1$,

$$\begin{aligned} \int_0^t dr \frac{dV}{dx}(X_r) &= \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) - \chi(\zeta_{\mathbf{N}_t} = 0) \int_t^{\tau_{\mathbf{N}_t+1}} dr \frac{dV}{dx}(X_r) \\ &\quad - \chi(\zeta_{\mathbf{N}_{t+1}} = 0) \int_{\tau_{\mathbf{N}_t+1}}^{\tau_{\mathbf{N}_{t+1}+2}} dr \frac{dV}{dx}(X_r) + \sum_{n=0}^{\mathbf{N}_t} \mathbf{C}_0^{(\lambda)}(\sigma_n) + \mathbf{m}_t + \mathbf{m}'_t, \end{aligned} \quad (4.40)$$

where \mathbf{m}_t and \mathbf{m}'_t correspond to odd and even contributions of the form

$$\begin{aligned}\mathbf{m}_t &= \sum_{n=1}^{\lfloor \frac{1}{2}\mathbf{N}_t - \frac{1}{2} \rfloor + 1} \left(\chi(\zeta_{2n} = 0) \int_{\tau_{2n}}^{\tau_{2n+1}} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(\sigma_{2n-1}) \right) \\ \mathbf{m}'_t &= \sum_{n=0}^{\lfloor \frac{1}{2}\mathbf{N}_t \rfloor} \left(\chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(\sigma_{2n}) \right).\end{aligned}$$

The processes $\mathbf{m}_t, \mathbf{m}'_t$ are not adapted to $\tilde{\mathcal{F}}_t$, since, for instance, \mathbf{m}'_t makes the jump

$$\chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(\sigma_{2n})$$

at time τ_{2n} , and the size of the jump depends on X_t up to time τ_{2n+2} . Let $\tilde{\mathcal{F}}_t''$ be the σ -algebra of all information before time τ_{n+2} , where $\tau_n \leq t < \tau_{n+1}$. The process \mathbf{m}'_t is a martingale with respect to $\tilde{\mathcal{F}}_t''$. To see this let us consider a time $t < \tau_{2n-1}$, then the following equalities hold:

$$\begin{aligned}\tilde{\mathbb{E}}_{\tilde{\nu}} \left[\chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_t'' \right] &= \tilde{\mathbb{E}}_{\tilde{\nu}} \left[\tilde{\mathbb{E}}_{\tilde{\nu}} \left[\chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_{\tau_{2n}}^- \right] \mid \tilde{\mathcal{F}}_t'' \right] \\ &= \tilde{\mathbb{E}}_{\tilde{\nu}} \left[\tilde{\mathbb{E}}_{\tilde{\delta}_{\sigma_{2n}}} \left[\chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) \mid \tilde{\mathcal{F}}_t'' \right] \right] \\ &= \tilde{\mathbb{E}}_{\tilde{\nu}} \left[\mathbf{C}_0^{(\lambda)}(\sigma_{2n}) \mid \tilde{\mathcal{F}}_t'' \right].\end{aligned}\tag{4.41}$$

The nested conditional expectation on the first line uses that $\tilde{\mathcal{F}}_t'' \subseteq \tilde{\mathcal{F}}_{\tau_{2n}}^-$, and the third equality is by definition of $\mathbf{C}_0^{(\lambda)}$. The second equality uses that $\zeta_{2n} = 1$ or $\zeta_{2n} = 0$ with probabilities respectively of $h(\sigma_{2n})$ and $1 - h(\sigma_{2n})$ given $\tilde{\mathcal{F}}_{\tau_{2n}}^-$ by Part 3 of Proposition 4.3, the strong Markov property starting from the time τ_{2n} , and the definition of the distribution $\tilde{\delta}_s$ (which is the splitting of the δ -distribution at $s \in S$). The predictable quadratic variation $\langle \mathbf{m}'_t \rangle$ for the martingale \mathbf{m}'_t has the form

$$\langle \mathbf{m}' \rangle_t = \sum_{n=0}^{\lfloor \frac{1}{2}\mathbf{N}_t \rfloor} \mathbf{C}_1^{(\lambda)}(\sigma_{2n}).\tag{4.42}$$

The analogous statements hold for \mathbf{m}_t .

By the triangle inequality,

$$\begin{aligned}\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t \leq R_1} \left| \int_0^t dr \frac{dV}{dx}(X_r) \right|^{2m} \right]^{\frac{1}{2m}} &\leq 6 \sup_x \left| \frac{dV}{dx}(x) \right| \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} (\tau_{\mathbf{N}_t+1} - \tau_{\mathbf{N}_t})^{2m} \right]^{\frac{1}{2m}} \\ &+ \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} \left(\sum_{n=0}^{\mathbf{N}_t} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right)^{2m} \right]^{\frac{1}{2m}} + \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} |\mathbf{m}_t|^{2m} \right]^{\frac{1}{2m}} + \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} |\mathbf{m}'_t|^{2m} \right]^{\frac{1}{2m}},\end{aligned}\tag{4.43}$$

where we have bounded each of the first three terms on the right side of (4.40) by the supremum of $|\frac{dV}{dx}(x)|$ multiplied by the longest interval $\tau_{n+1} - \tau_n$ over $n \leq \tilde{n}_1$. We used a factor of 6 rather than 3, since there is one term for which the interval $[\tau_{2n+1}, \tau_{2n+2}]$ will have $\tau_{2n+1} \in [R_1, R_2]$

rather than $< R_1$. We thus double the bound, since we can apply the Markov property for stating from time R_1 . We now look at the terms on the right side one-by-one.

For the first term on the right side of (4.43),

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} (\tau_{\mathbf{N}_{t+1}} - \tau_{\mathbf{N}_t})^{2m} \right] &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq n \leq \tilde{n}_1} (\tau_{n+1} - \tau_n)^{2m} \right] \\ &\leq c^{2m} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\mathbb{E} \left[\sup_{0 \leq n \leq \tilde{n}_1} \mathbf{e}_n^{2m} \mid \tilde{n}_1 \right] \right] \leq c' + c' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[(\log(\tilde{n}_1))^{2m} \right], \end{aligned} \quad (4.44)$$

where \mathbf{e}_n are i.i.d. mean-1 exponential random variables independent of everything else. The $c > 0$ in the first inequality is from Lemma 4.6 and replacing the $(\tau_{n+1} - \tau_n)$'s with the \mathbf{e}_n 's. The $c' > 0$ for the second inequality exists by an elementary analysis of $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq n \leq N} \mathbf{e}_n^{2m} \right]$ for $N > 0$ and independent exponential random variables \mathbf{e}_n with mean one. The value $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[(\log(\tilde{n}_1))^{2m} \right]$ is finite by Proposition 4.7, since the fractional moments $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1^\alpha]$ are finite for $0 < \alpha < \frac{1}{2}$.

For the second term on the right side of (4.43), obviously

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} \left(\sum_{n=0}^{\mathbf{N}_t} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right)^{2m} \right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\sum_{n=0}^{\tilde{n}_1} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right)^{2m} \right],$$

since $\mathbf{N}_t = \tilde{n}_1$ for $t \in [R_{n-1}, R_n)$. By Lemma 4.13, $g_\lambda = \mathbf{C}_0^{(\lambda)}$ has the inequality $g_\lambda(x, p) \leq C(\frac{1}{1+p^2} \vee \lambda)$ for $\lambda < 1$. Hence by Lemma 4.10, the above sum is bounded independently of $\lambda < 1$.

The last two terms on the right side of (4.43) are similar, so we will only treat the last. By Rosenthal's inequality (see e.g. [11], Lemma 2.1), there is a C'' such that

$$\begin{aligned} &\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq t < R_1} |\mathbf{m}'_t|^{2m} \right] \\ &\leq C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\langle \mathbf{m}' \rangle_{R_1}^m] + C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq n \leq \lfloor \frac{\tilde{n}_1}{2} \rfloor} \left| \chi(\zeta_{2n+1} = 0) \int_{\tau_{2n+1}}^{\tau_{2n+2}} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(\sigma_{2n}) \right|^{2m} \right] \\ &\leq C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\langle \mathbf{m}' \rangle_{R_1}^m] + C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \left| \chi(\zeta_{n+1} = 0) \int_{\tau_{n+1}}^{\tau_{n+2}} dr \frac{dV}{dx}(X_r) - \mathbf{C}_0^{(\lambda)}(\sigma_n) \right|^{4m} \right]^{\frac{1}{2}} \\ &\leq C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\sum_{n=0}^{\tilde{n}_1+1} \mathbf{C}_1^{(\lambda)}(\sigma_n) \right)^m \right] + C'' \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} \mathbf{C}_{2m}^{(\lambda)}(\sigma_n) \right]^{\frac{1}{2}}. \end{aligned}$$

For the second inequality, we have used the standard technique to bound the supremum in the second term with $(\sup_n a_n)^2 \leq \sum_n a_n^2$ and Jensen's inequality, and we have also included the odd terms. The first term in the third inequality is bounded with the equality (4.42) and by including the odd terms. For the second term in the third equality is by inserting a nested conditional expectation with respect to $\tilde{\mathcal{F}}_{\tau_{n-1}}^-$ for the n th term in sum and applying the argument in (4.41). To bound both terms on the right side, we apply Lemma 4.13 and Lemma 4.10 as above.

Part (2): Again, we start by writing $\int_0^{R_1} dr \frac{dV}{dx}(X_r)$ as in (4.40), and use the triangle inequality

to get

$$\begin{aligned} \tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_0^{R_1} dr \frac{dV}{dx}(X_r) \right| \right] &\leq \tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right| \right] + \tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_t^{\tau_{\tilde{n}_1+1}} dr \frac{dV}{dx}(X_r) \right| \right] \\ &+ \tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_{\tau_{\tilde{n}_1+1}}^{\tau_{\tilde{n}_1+2}} dr \frac{dV}{dx}(X_r) \right| \right] + \tilde{\mathbb{E}}_s^{(\lambda)} \left[\sum_{n=0}^{\tilde{n}_1} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right] + \tilde{\mathbb{E}}_s^{(\lambda)} [|\mathbf{m}_{R_1}|^2]^{\frac{1}{2}} + \tilde{\mathbb{E}}_s^{(\lambda)} [|\mathbf{m}'_{R_1}|^2]^{\frac{1}{2}}, \quad (4.45) \end{aligned}$$

where we have also applied Jensen's inequality to the last two terms. The first three terms on the right side are bounded by $c \sup_x |\frac{dV}{dx}(x)|$, where $c > 0$ is from Lemma 4.6. This follows for the first term, for instance, since

$$\tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right| \right] \leq \sup_x \left| \frac{dV}{dx}(x) \right| \tilde{\mathbb{E}}_s^{(\lambda)} [\tau_1] \leq c \sup_x \left| \frac{dV}{dx}(x) \right|,$$

where the second inequality is by Lemma 4.6.

Since \mathbf{m}'_t is a martingale, we have the first equality below

$$\tilde{\mathbb{E}}_s^{(\lambda)} [|\mathbf{m}'_{R_1}|^2] = \tilde{\mathbb{E}}_s^{(\lambda)} [\langle \mathbf{m}' \rangle_{R_1}] = \tilde{\mathbb{E}}_s^{(\lambda)} \left[\sum_{n=1}^{\lfloor \frac{1}{2} \tilde{n}_1 \rfloor} \mathbf{C}_1^{(\lambda)}(\sigma_{2n}) \right] \leq \tilde{\mathbb{E}}_s^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} \mathbf{C}_1^{(\lambda)}(\sigma_n) \right].$$

A similar calculation holds for the term $\tilde{\mathbb{E}}_s^{(\lambda)} [|\mathbf{m}_{R_1}|^2]$. With the above, then

$$\begin{aligned} \tilde{\mathbb{E}}_s^{(\lambda)} \left[\left| \int_0^{R_1} dr \frac{dV}{dx}(X_r) \right| \right] &\leq 3c \sup_x \left| \frac{dV}{dx}(x) \right| + \tilde{\mathbb{E}}_s^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right] + 2\tilde{\mathbb{E}}_s^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} \mathbf{C}_1^{(\lambda)}(\sigma_n) \right]^{\frac{1}{2}} \\ &\leq 3c \sup_x \left| \frac{dV}{dx}(x) \right| + b \mathbb{E}_{(x,p)}^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} |\mathbf{C}_0^{(\lambda)}(\sigma_n)| \right] + b^{\frac{1}{2}} \mathbb{E}_{(x,p)}^{(\lambda)} \left[\sum_{n=1}^{\tilde{n}_1} \mathbf{C}_1^{(\lambda)}(\sigma_n) \right]^{\frac{1}{2}} \\ &= 3c \sup_x \left| \frac{dV}{dx}(x) \right| + b(W^{(\lambda)} \mathbf{C}_0^{(\lambda)})(x, p) + b^{\frac{1}{2}} \left((W^{(\lambda)} \mathbf{C}_1^{(\lambda)})(x, p) \right)^{\frac{1}{2}}, \end{aligned}$$

for $b = \frac{U}{U-\mathbf{u}} \vee \frac{U}{\mathbf{u}}$ and the map $W^{(\lambda)} : L^\infty(\Sigma) \rightarrow L^\infty(\Sigma)$ was defined in Section 4.3. The second inequality uses that $\mathbb{E}_s = (1 - h(s))\tilde{\mathbb{E}}_{(s,0)} + h(s)\tilde{\mathbb{E}}_{(s,1)}$. With Part (1) of Lemma 4.13 for $g_\lambda(s) = \mathbf{C}_0^{(\lambda)}(s)$ and $g_\lambda(s) = \mathbf{C}_1^{(\lambda)}(s)$ and Lemma 4.9, we obtain a logarithmic bound in $|p|$ for the right side.

Part (3): We can split the expectation into two terms as

$$\begin{aligned} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr \frac{dV}{dx}(X_r) \int_r^{R_2} dr' \frac{dV}{dx}(X_{r'}) \right] \\ = \frac{1}{2} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\int_0^{R_1} dr \frac{dV}{dx}(X_r) \right)^2 \right] + \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_{R'_1}^{R_1} dr \frac{dV}{dx}(X_r) \int_r^{R_2} dr' \frac{dV}{dx}(X_{r'}) \right] \\ = \frac{1}{2} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right)^2 \right] + \tilde{\mathbb{E}}_{\tilde{\nu}'}^{(\lambda)} \left[\int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \int_r^{R_1} dr' \frac{dV}{dx}(X_{r'}) \right], \end{aligned}$$

where $\tilde{\nu}'$ is the distribution for $\tilde{S}_{R'_1} \in \Sigma \times 1$ when \tilde{S}_0 is distributed as $\tilde{\nu}$ and has the density

$$\tilde{\nu}'(s, z) = \chi(z=1) \frac{\Psi_{\infty, \lambda}(s) h(s)}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)}.$$

The second equality above is by the strong Markov property for the split resolvent chain corresponding to the time $R'_1 = \tau_{\tilde{n}_1}$. The two terms on the right are handled similarly, and we will treat the first.

As in Part (2), we can rewrite the argument of the expectation as a function of the first component of the chain $\tilde{\sigma}_n = (\sigma_n, \zeta_n)$ up to $n = \tilde{n}_1$ rather than the process (\tilde{S}_t) , since

$$\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right)^2 \right] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[f^{(\lambda)}(\sigma_0, \dots, \sigma_{\tilde{n}_1}) \right] = \int_{\Upsilon} d\psi_{\lambda}(\xi) f^{(\lambda)}(\xi),$$

where Υ is the set of all trajectories $\xi = (\tilde{\sigma}_0, \dots, \tilde{\sigma}_{\tilde{n}_1})$ starting in the distribution $\tilde{\nu}$ and ending at the first return to the set $\Sigma \times 1$, ψ_{λ} is the measure induced on Υ , and $f^{(\lambda)}$ is

$$f^{(\lambda)}(\sigma_0, \dots, \sigma_{\tilde{n}_1}) = \sum_{n=0}^{\tilde{n}_1-1} g_2^{(\lambda)}(\sigma_n, \sigma_{n+1}) + 2 \sum_{n=0}^{\tilde{n}_1-1} g_1^{(\lambda)}(\sigma_n, \sigma_{n+1}) \sum_{n+1=m}^{\tilde{n}_1-1} g_1^{(\lambda)}(\sigma_m, \sigma_{m+1}),$$

for $g_m^{(\lambda)} : \Sigma^2 \rightarrow \mathbb{R}$ defined as

$$g_m^{(\lambda)}(s_1, s_2) = \mathbb{E}_{s_1}^{(\lambda)} \left[\left(\int_0^{\tau_1} dr \frac{dV}{dx}(X_r) \right)^m \middle| S_{\tau_1} = s_2 \right].$$

Note that $f^{(\lambda)}$ is positive, since it is the conditional expectation of $\left| \int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right|^2$ given the trajectory ξ . Let Υ_{λ} be the event that $\tilde{n}_1 \leq \lambda^{-\frac{1}{9}}$ and $|P_{\tau_n}| \leq \lambda^{-\frac{1}{4}}$ for $n \leq \tilde{n}_1$. By the triangle inequality,

$$\begin{aligned} \left| \int_{\Upsilon} d\psi_{\lambda}(\xi) f^{(\lambda)}(\xi) - \int_{\Upsilon} d\psi_0(\xi) f^{(0)}(\xi) \right| &\leq \left| \int_{\Upsilon_{\lambda}} d\psi_{\lambda}(\xi) f^{(\lambda)}(\xi) - \int_{\Upsilon_{\lambda}} d\psi_0(\xi) f^{(0)}(\xi) \right| \\ &\quad + \int_{\Upsilon_{\lambda}^c} d\psi_{\lambda}(\xi) f^{(\lambda)}(\xi) + \int_{\Upsilon_{\lambda}^c} d\psi_0(\xi) f^{(0)}(\xi). \end{aligned} \quad (4.46)$$

The last two terms on the right side are similar, and the first of them is smaller than

$$\begin{aligned} \int_{\Upsilon_{\lambda}^c} d\psi_{\lambda}(\xi) f^{(\lambda)}(\xi) &= \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\chi(\Upsilon_{\lambda}^c) \left(\int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right)^2 \right] \\ &\leq \tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} [\Upsilon_{\lambda}^c]^{\frac{1}{2}} \sup_{\lambda < 1} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\left(\int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right)^4 \right]^{\frac{1}{2}}, \end{aligned}$$

where we have applied the Cauchy-Schwarz inequality. By Part (3), the rightmost term is finite. We will show that $\tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} [\Upsilon_{\lambda}^c] = O(\lambda^{\frac{1}{20}})$. Since $|P_r| \leq \mathbf{Q}_r = 2^{\frac{1}{2}} H_r^{\frac{1}{2}}$, we have that

$$\tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} [\Upsilon_{\lambda}^c] \leq \tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1 > \lambda^{-\frac{1}{9}}] + \tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{0 \leq m \leq \lambda^{-\frac{1}{9}}} \mathbf{Q}_{\tau_m} > \lambda^{-\frac{1}{4}} \right]. \quad (4.47)$$

For the first term on the right side of (4.47), Chebyshev's inequality gives

$$\tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1 \geq \lambda^{-\frac{1}{9}}] \leq \lambda^{\frac{\alpha}{9}} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [\tilde{n}_1^{\alpha}] \leq C_{\alpha} \lambda^{\frac{\alpha}{9}},$$

where $0 \leq \alpha < \frac{1}{2}$ and the C_α of the second inequality is from Proposition 4.7. We can choose $\alpha = \frac{9}{20}$. For the second term on the right side of (4.47),

$$\begin{aligned} \tilde{\mathbb{P}}_\nu^{(\lambda)} \left[\sup_{0 \leq m \leq \lambda^{-\frac{1}{9}}} \mathbf{Q}_{\tau_m} > \lambda^{-\frac{1}{4}} \right] &= \mathbb{P}_\nu^{(\lambda)} \left[\sup_{0 \leq m \leq \lambda^{-\frac{1}{9}}} \mathbf{Q}_{\tau_m} > \lambda^{-\frac{1}{4}} \right] \leq \mathbb{P}_\nu^{(\lambda)} \left[\sup_{0 \leq r \leq 2\lambda^{-\frac{1}{9}}} \mathbf{Q}_r > \lambda^{-\frac{1}{4}} \right] + O(\lambda) \\ &\leq \lambda^{\frac{1}{2}} \mathbb{E}_\nu^{(\lambda)} \left[\sup_{0 \leq r \leq 2\lambda^{-\frac{1}{9}}} \mathbf{Q}_r^2 \right] + O(\lambda), \end{aligned}$$

where $T_\lambda = \tau_m$ for $m = \lfloor \lambda^{-\frac{1}{9}} \rfloor$. The first equality holds, since the argument of the probability is well-defined in the original dynamics. Since the times τ_m are mean- m exponential, the probability that $\tau_m > 2\lambda^{-\frac{1}{9}}$ for $m = \lfloor \lambda^{-\frac{1}{9}} \rfloor$ is super-polynomially small. There is a $C > 0$ such that for all $\lambda < 1$,

$$\mathbb{E}_\nu^{(\lambda)} \left[\sup_{0 \leq r \leq T} \mathbf{Q}_r^2 \right] \leq CT$$

by similar techniques as for in the proof in the Lemma 2.2. Plugging in with $T = 2\lambda^{-\frac{1}{9}}$ implies that the rightmost term in (4.47) is $O(\lambda^{\frac{7}{18}})$ and thus $\tilde{\mathbb{P}}_\nu^{(\lambda)}[\Upsilon_\lambda^c] = O(\lambda^{\frac{1}{20}})$.

For the first term on the right side of (4.46),

$$\begin{aligned} &\left| \int_{\Upsilon_\lambda} d\psi_\lambda(\xi) f^{(\lambda)}(\xi) - \int_{\Upsilon_\lambda} d\psi_0(\xi) f^{(0)}(\xi) \right| \\ &\leq \left| \int_{\Upsilon_\lambda} (d\psi_\lambda(\xi) - d\psi_0(\xi)) f^{(0)}(\xi) \right| + \int_{\Upsilon_\lambda} d\psi_\lambda(\xi) |f^{(\lambda)}(\xi) - f^{(0)}(\xi)|. \end{aligned} \quad (4.48)$$

Let $\psi_\lambda^{(n)}$ be the measure over Υ or Υ_λ in which the first n steps are made with transition measure $\tilde{\mathcal{T}}_0$ and the remaining are with respect to $\tilde{\mathcal{T}}_\lambda$. Thus $\psi_\lambda^{(0)} = \psi_\lambda$, and $\psi_\lambda^{(N)} = \psi_0$ for $N \geq \lfloor \lambda^{-\frac{1}{9}} \rfloor$ in the event Υ_λ . For $\xi \in \Upsilon_\lambda$, then applying Part (1) of Lemma 4.8

$$|d\psi_\lambda^{(n+1)}(\xi) - d\psi_\lambda^{(n)}(\xi)| \leq \chi(\tilde{n}_1 > n) bc\lambda^{\frac{1}{4}} d\psi_\lambda^{(n)}(\xi).$$

Employing this inductively, then we have the two inequalities

$$\begin{aligned} |d\psi_0^{(n+1)}(\xi) - d\psi_\lambda^{(n)}(\xi)| &\leq \chi(\tilde{n}_1 > n) bc\lambda^{\frac{1}{4}} e^{bc(\tilde{n}_1 - n)\lambda^{\frac{1}{4}}} d\psi_0(\xi) \\ &\leq \chi(\tilde{n}_1 > n) bc\lambda^{\frac{1}{4}} e^{bc\lambda^{\frac{1}{8}}} d\psi_0(\xi). \end{aligned} \quad (4.49)$$

By adding and subtracting the terms $d\psi_\lambda^{(n)}(\xi)$, $0 < n < \lfloor \lambda^{-\frac{1}{9}} \rfloor$ as a telescoping sum and applying the triangle inequality and (4.49),

$$|d\psi_0(\xi) - d\psi_\lambda(\xi)| \leq \sum_{n=0}^{\lfloor \lambda^{-\frac{1}{9}} \rfloor - 1} |d\psi_0^{(n+1)}(\xi) - d\psi_\lambda^{(n)}(\xi)| \leq bc\lambda^{\frac{1}{8}} e^{bc\lambda^{\frac{1}{8}}} d\psi_0(\xi). \quad (4.50)$$

Thus with (4.50), the first term on the right side of (4.48) is bounded by

$$\begin{aligned} &\left| \int_{\Upsilon_\lambda} (d\psi_\lambda(\xi) - d\psi_0(\xi)) f^{(0)}(\xi) \right| \leq \int_{\Upsilon_\lambda} |d\psi_\lambda(\xi) - d\psi_0(\xi)| f^{(0)}(\xi) \\ &\leq bc\lambda^{\frac{1}{8}} e^{bc\lambda^{\frac{1}{8}}} \int_{\Upsilon} d\psi_0(\xi) f^{(0)}(\xi) = bc\lambda^{\frac{1}{8}} e^{bc\lambda^{\frac{1}{8}}} \mathbb{E}_\nu^{(0)} \left[\left(\int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right)^2 \right] = O(\lambda^{\frac{1}{8}}), \end{aligned} \quad (4.51)$$

where the third inequality follows by extending the domain of integration.

For the second term on the right side of (4.48),

$$\begin{aligned}
& \int_{\Upsilon_\lambda} d\psi_\lambda(\xi) |f^{(\lambda)}(\xi) - f^{(0)}(\xi)| \leq e^{bc\lambda^{\frac{1}{8}}} \int_{\Upsilon_\lambda} d\psi_0(\xi) |f^{(\lambda)}(\xi) - f^{(0)}(\xi)| \\
& \leq 2e^{bc\lambda^{\frac{1}{8}}} \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} |g_2^{(\lambda)}(\sigma_n, \sigma_{n+1}) - g_2^{(0)}(\sigma_n, \sigma_{n+1})| \\
& + 4e^{bc\lambda^{\frac{1}{8}}} \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} |g_1^{(\lambda)}(\sigma_n, \sigma_{n+1}) - g_1^{(0)}(\sigma_n, \sigma_{n+1})| \sum_{m=n+1}^{\tilde{n}_1-1} |g_1^{(0)}(\sigma_m, \sigma_{m+1})| \\
& + 4e^{bc\lambda^{\frac{1}{8}}} \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} |g_1^{(\lambda)}(\sigma_n, \sigma_{n+1})| \sum_{m=n+1}^{\tilde{n}_1-1} |g_1^{(0)}(\sigma_m, \sigma_{m+1}) - g_1^{(\lambda)}(\sigma_m, \sigma_{m+1})|.
\end{aligned}$$

The three terms on the right side are handled similarly, so we will address the second.

$$\begin{aligned}
& \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} |g_1^{(\lambda)}(\sigma_n, \sigma_{n+1}) - g_1^{(0)}(\sigma_n, \sigma_{n+1})| \sum_{m=n+1}^{\tilde{n}_1-1} |g_1^{(0)}(\sigma_m, \sigma_{m+1})| \\
& \leq \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} w_1^{(\lambda)}(\sigma_n) \sum_{m=n+1}^{\tilde{n}_1-1} w_2^{(\lambda)}(\sigma_m) \leq \lambda^{-\frac{1}{8}} b \sup_a \left| \frac{dV}{dx}(a) \right| \int_{\Upsilon_\lambda} d\psi_0(\xi) \sum_{n=0}^{\tilde{n}_1-1} w_1^{(\lambda)}(\sigma_n) \\
& \leq \lambda^{-\frac{1}{4}} b \left(\sup_a \left| \frac{dV}{dx}(a) \right| \right) \sup_{|p'| \leq \lambda^{-\frac{1}{4}}} w_1^{(\lambda)}(x', p', z') = O(\lambda^{\frac{1}{4}}), \tag{4.52}
\end{aligned}$$

where $b = \frac{U-\mathbf{u}}{\mathbf{u}} \vee \frac{U}{\mathbf{u}}$ and $w_1, w_2 : \tilde{\Sigma} \rightarrow \mathbb{R}^+$ are defined as

$$w_1^{(\lambda)}(\tilde{s}) = \int_{\tilde{\Sigma}} \tilde{\mathcal{T}}^{(0)}(\tilde{s}, d\tilde{s}') |g_1^{(\lambda)}(s, s') - g_1^{(0)}(s, s')| \quad \text{and} \quad w_2^{(\lambda)}(\tilde{s}) = \int_{\tilde{\Sigma}} \tilde{\mathcal{T}}^{(0)}(\tilde{s}, d\tilde{s}') |g_1^{(0)}(s, s')|.$$

For the second inequality of (4.52), we have used that $\tilde{n}_1 \leq \lambda^{-\frac{1}{8}}$ for $\xi \in \Upsilon_\lambda$ and

$$w_2^{(\lambda)}(\tilde{s}) \leq \tilde{\mathbb{E}}_{\tilde{s}} \left[\left| \int_0^\tau dr \frac{dV}{dx}(X_r) \right| \right] \leq \left(\sup_x \left| \frac{dV}{dx}(x) \right| \right) \tilde{\mathbb{E}}_{\tilde{s}}[\tau] \leq b \sup_x \left| \frac{dV}{dx}(x) \right|,$$

and the last inequality is by Lemma 4.6. Moreover,

$$\begin{aligned}
w_1^{(\lambda)}(\tilde{s}) &= \tilde{\mathbb{E}}_{\tilde{s}}^{(\lambda)} [|g_1^{(\lambda)}(S_0, S_\tau) - g_1^{(0)}(S_0, S_\tau)|] \\
&\leq b \mathbb{E}_s^{(\lambda)} [|g_1^{(\lambda)}(S_0, S_\tau) - g_1^{(0)}(S_0, S_\tau)|] = b \int_{\Sigma} ds' \mathcal{T}^{(0)}(s, s') |g_1^{(\lambda)}(s, s') - g_1^{(0)}(s, s')|.
\end{aligned}$$

The first inequality follows by the usual relation $\mathbb{E}_s^{(\lambda)} = (1 - h(s))\tilde{\mathbb{E}}_{(s,0)}^{(\lambda)} + h(s)\tilde{\mathbb{E}}_{(s,1)}^{(\lambda)}$. By Part (2) of Lemma 4.8, the supremum of the right side all $s = (x, p)$ with $|p| \leq \lambda^{-\frac{1}{4}}$ is $O(\lambda^{\frac{1}{2}})$, and thus we get the order equality in (4.52). \square

4.5 Proof that cumulative forcing vanishes

Now we nearly have the material required to show the convergence $\sup_{0 \leq t \leq T} |\lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}}| \implies 0$, and more precisely, is $O(\lambda^{\frac{1}{4}})$.

Part of our argument will rely on an analysis of certain processes related to the energy H_r . Recall that the processes \mathbf{Q}_t , \mathbf{M}_t , and \mathbf{A}_r are such that

$$\mathbf{Q}_t = (2H_t)^{\frac{1}{2}} = \mathbf{Q}_0 + \mathbf{M}_t + \mathbf{A}_t.$$

The processes \mathbf{M}_t and \mathbf{A}_t are the martingale and predictable parts in the semimartingale decomposition of \mathbf{Q}_t , and both are initially zero. Let $\mathcal{A}_\lambda, \mathcal{A}_\lambda^\pm, \mathcal{V}_\lambda : \mathbb{T} \times \mathbb{R} \rightarrow \mathbb{R}$ be defined as in Section 2. The predictable component has the form $\mathbf{A}_t = \int_0^t dr \mathcal{A}_\lambda(X_r, P_r)$. We also define $\mathbf{A}_t^\pm = \int_0^t dr \mathcal{A}_\lambda^\pm(X_r, P_r)$. The martingale \mathbf{M}_t has predictable quadratic variation $\langle \mathbf{M} \rangle_t = \int_0^t dr \mathcal{V}_\lambda(X_r, P_r)$. The proposition below is an extension of Proposition 2.1 and contains some more bounds regarding the functions \mathcal{A}_λ^\pm , \mathcal{V}_λ , and $\mathcal{K}_{\lambda,n}$.

Proposition 4.15. *There exists $C, c > 0$ such that for λ small enough,*

1. *For $\lambda^{-\frac{3}{8}} \leq |p| \leq \lambda^{-\frac{3}{4}}$,*

$$\left| \mathcal{A}_\lambda^-(x, p) - \frac{1}{2} \lambda |p| \right| \leq C \lambda^{\frac{5}{4}} |p|, \quad \left| \mathcal{V}_\lambda(x, p) - 1 \right| \leq C \lambda^{\frac{1}{2}}, \quad \text{and} \quad \left| 2\mathcal{K}_{\lambda,2}(x, p) - 1 \right| \leq C \lambda^{\frac{1}{2}}.$$

2. *For all $(x, p) \in \Sigma$, $\mathcal{A}_\lambda^-(x, p) \leq |\mathcal{D}_\lambda(p)|$. In particular, for $|p| \leq \lambda^{-1}$, one has $\mathcal{A}_\lambda^-(x, p) \leq C \lambda |p|$.*

3. *For all $(x, p) \in \Sigma$ with $|p| \geq \lambda^{-1}$, $\frac{\mathcal{A}_\lambda^-(x, p)}{\mathcal{E}_\lambda(p)} \geq c > 0$.*

4. *For all $(x, p) \in \Sigma$, $\left| \frac{\mathcal{A}_\lambda(x, p)}{\mathcal{E}_\lambda(p)} + \frac{2\lambda|p|}{1+\lambda} \right| \leq C$.*

5. *As $\lambda \rightarrow 0$, we have*

$$\int_\Sigma ds \mathcal{A}_\lambda^+(s) = 1 + O(\lambda).$$

The following proposition lists two easy facts that will be referred to often.

Proposition 4.16. *For the split statistics, $\tilde{N}_t - \sum_{n=1}^{\mathbf{N}_t} h(S_{\tau_n})$ is a martingale with respect to the filtration $\tilde{\mathcal{F}}_t$. For the original statistics, $\sum_{n=1}^{\mathbf{N}_t} h(S_{\tau_n}) - \int_0^t dr h(S_r)$ is a martingale with respect to \mathcal{F}_t . In particular,*

$$\tilde{\mathbb{E}}^{(\lambda)}[\tilde{N}_t] = \mathbb{E}^{(\lambda)}\left[\int_0^t dr h(S_r)\right].$$

Proof. The difference

$$\tilde{N}_t - \sum_{n=1}^{\mathbf{N}_t} h(S_{\tau_n}) = \sum_{n=1}^{\mathbf{N}_t} (\chi(z_{\tau_n} = 1) - h(S_{\tau_n}))$$

is a martingale by Part (3) of Proposition 4.3. The difference $\sum_{n=1}^{\mathbf{N}_t} h(S_{\tau_n}) - \int_0^t dr h(S_r)$ is a martingale according to the pre-split law, since the contributions $h(S_{\tau_n})$ occur with Poisson rate 1. □

Define $L_t = \mathbf{u}^{-1} \int_0^t dr h(X_r, P_r)$. The following lemma states that rescaled versions of L_t and \mathbf{A}_t^+ are close, and gives a bound for the rescaled version of L_t . $\langle \tilde{M} \rangle_t$ is the predictable quadratic variation of the martingale \tilde{M} from Lemma 4.5.

Lemma 4.17. *As $\lambda \rightarrow 0$,*

1.

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} \left| \lambda^{\frac{1}{2}} L_{\frac{t}{\lambda}} - \lambda^{\frac{1}{2}} \mathbf{A}_{\frac{t}{\lambda}}^+ \right| \right] = O(\lambda^{\frac{1}{4}}).$$

Moreover, there is a $C > 0$ such that for $\lambda < 1$, then

$$\mathbb{E}^{(\lambda)} \left[\lambda^{\frac{1}{2}} L_{\frac{T}{\lambda}} \right] \leq C.$$

2.

$$\tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq T} \left| \lambda^{\frac{1}{2}} \langle \tilde{M} \rangle_{\frac{t}{\lambda}} - \lambda^{\frac{1}{2}} v_{\lambda} \tilde{N}_{\frac{t}{\lambda}} \right| \right] = O(\lambda^{\frac{1}{4}}).$$

Also, for any $t \geq 0$, the expectations are equal $\tilde{\mathbb{E}}^{(\lambda)} [\langle \tilde{M} \rangle_t] = v_{\lambda} \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_t]$.

Proof.

Part (1): First, we will show that there is a $C > 0$ such that for all $\lambda < 1$

$$\tilde{\mathbb{E}}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \tilde{N}_{\frac{T}{\lambda}} \right] \leq C.$$

Define the constant

$$a_{\lambda} := \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\int_0^{R_1} dr \mathcal{A}_{\lambda}^+(S_r) \right] = \frac{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) \mathcal{A}_{\lambda}^+(s)}{\int_{\Sigma} ds \Psi_{\infty, \lambda}(s) h(s)} = \frac{\int_{\Sigma} dx dp e^{-\lambda(\frac{1}{2}p^2 + V(x))} \mathcal{A}_{\lambda}^+(x, p)}{\int_{\Sigma} dx dp e^{-\lambda(\frac{1}{2}p^2 + V(x))} h(x, p)},$$

where the second equality holds by Part (1) of Proposition 4.11. By Part (5) of Proposition 4.15, the compact support for h , and the fact $\int_{\Sigma} ds h(s)$, we have that $a_{\lambda} = \mathbf{u}^{-1} + O(\lambda)$. The expectation of $\lambda^{\frac{1}{2}} \tilde{N}_{\frac{T}{\lambda}}$ is bounded by

$$\tilde{\mathbb{E}}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \tilde{N}_{\frac{T}{\lambda}} \right] = \lambda^{\frac{1}{2}} + \tilde{\mathbb{E}}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \sum_{n=1}^{\tilde{N}_{\frac{T}{\lambda}} - 1} a_{\lambda}^{-1} \int_{R_n}^{R_{n+1}} dr \mathcal{A}_{\lambda}^+(S_r) \right] \quad (4.53)$$

$$\leq \lambda^{\frac{1}{2}} + \frac{2}{\mathbf{u}} \tilde{\mathbb{E}}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \int_0^{\frac{T}{\lambda}} dr \mathcal{A}_{\lambda}^+(S_r) \right] = \lambda^{\frac{1}{2}} + \frac{2}{\mathbf{u}} \mathbb{E}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \mathbf{A}_{\frac{T}{\lambda}}^+ \right]. \quad (4.54)$$

The first equality uses the Markov property and the definition of a_{λ} , and the second equality merely goes back to the pre-split statistics. The inequality is for λ small enough such that $a_{\lambda} \geq \frac{1}{2} \mathbf{u}$.

Since $\mathbf{A}_t^+ = \mathbf{Q}_t - \mathbf{Q}_0 - \mathbf{M}_t + \mathbf{A}_t^-$, then by the triangle inequality,

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \mathbf{A}_{\frac{T}{\lambda}}^+ \right] &\leq 2\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \lambda^{\frac{1}{2}} \mathbf{Q}_t \right] + \mathbb{E}^{(\lambda)} \left[|\lambda^{\frac{1}{2}} \mathbf{M}_{\frac{T}{\lambda}}| \right] + \mathbb{E}^{(\lambda)} \left[|\lambda^{\frac{1}{2}} \mathbf{A}_{\frac{T}{\lambda}}^-| \right] \\ &\leq 2\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \lambda^{\frac{1}{2}} \mathbf{Q}_t \right] + \mathbb{E}^{(\lambda)} \left[\lambda \int_0^{\frac{T}{\lambda}} dr \mathcal{V}_{\lambda}(S_r) \right]^{\frac{1}{2}} + \mathbb{E}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \int_0^{\frac{T}{\lambda}} |\mathcal{A}_{\lambda}^-(S_r)| \right] \\ &\leq 2\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \lambda^{\frac{1}{2}} \mathbf{Q}_t \right] + C_1^{\frac{1}{2}} \mathbb{E}^{(\lambda)} \left[\lambda \int_0^{\frac{T}{\lambda}} dr (1 + \lambda \mathbf{Q}_r) \right]^{\frac{1}{2}} + C_2 \mathbb{E}^{(\lambda)} \left[\lambda^{\frac{1}{2}} \int_0^{\frac{T}{\lambda}} dr \lambda \mathbf{Q}_r \right] \end{aligned} \quad (4.55)$$

For the second inequality, the second term employs Jensen's inequality with the square function along with the fact that the martingale \mathbf{M}_t has bracket $\langle \mathbf{M} \rangle_t = \int_0^t dr \mathcal{V}_\lambda(S_r)$. The bound for the second term in the third equality is Part (3) of Proposition 2.1 and $\mathcal{E}_\lambda(P_r) \leq c(1 + \lambda|P_r|) \leq c(1 + \lambda\mathbf{Q}_r)$ for some $c > 0$ by Part (2) of Proposition 3.1. The third term in the third inequality is Part (3) of Proposition 4.15 and $|P_r| \leq \mathbf{Q}_r$. The right side of (4.55) is uniformly bounded in λ , since $\mathbb{E}^{(\lambda)}[\mathbf{Q}_r] \leq \mathbb{E}^{(\lambda)}[2H_r]^{\frac{1}{2}}$ by Jensen's, and $\sup_{0 \leq r \leq t} \mathbb{E}^{(\lambda)}[H_r] \leq C\lambda^{-1}$ for some $C > 0$ by Lemma 2.2.

It follows from the above that $\tilde{\mathbb{E}}^{(\lambda)}[\lambda^{\frac{1}{2}}\tilde{N}_{\frac{T}{\lambda}}]$ is bounded uniformly for small λ . Since

$$\mathbf{u}\tilde{\mathbb{E}}^{(\lambda)}[\tilde{N}_t] = \tilde{\mathbb{E}}^{(\lambda)}[L_t]$$

by Part (1) of Proposition 4.16, we also have that $\tilde{\mathbb{E}}^{(\lambda)}[\lambda^{\frac{1}{2}}L_{\frac{T}{\lambda}}]$ is uniformly bounded.

To show that $\lambda^{\frac{1}{2}}L_{\frac{T}{\lambda}}$ is close to $\lambda^{\frac{1}{2}}\mathbf{A}_{\frac{T}{\lambda}}^+$, we will consider the split dynamics.

$$\begin{aligned} \mathbb{E}^{(\lambda)}\left[\sup_{0 \leq t \leq T} \left|\lambda^{\frac{1}{2}}L_{\frac{t}{\lambda}} - \lambda^{\frac{1}{2}}\mathbf{A}_{\frac{t}{\lambda}}^+\right|\right] &= \tilde{\mathbb{E}}^{(\lambda)}\left[\sup_{0 \leq t \leq T} \left|\lambda^{\frac{1}{2}}L_{\frac{t}{\lambda}} - \lambda^{\frac{1}{2}}\mathbf{A}_{\frac{t}{\lambda}}^+\right|\right] \\ &\leq \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\left(\int_0^{R_1} dr + \sup_{0 \leq t \leq \frac{T}{\lambda}} \int_t^{R_{\tilde{N}_t+1}} dr\right) g_\lambda(S_r)\right] + \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left|\sum_{n=1}^{\tilde{N}_t} \int_{R_n}^{R_{n+1}} g'_\lambda(S_r)\right|\right], \end{aligned} \quad (4.56)$$

where R_n is the starting time of the n life cycle, \tilde{N}_t is the number of R'_n (not necessarily R_n) which have occurred up to time t , $g_\lambda = \mathcal{A}_\lambda^+ + \mathbf{u}^{-1}h$, and $g'_\lambda = \mathcal{A}_\lambda^+ - \mathbf{u}^{-1}h$.

The first term on the right side contains the boundary terms for the partition of the integrals over the interval $[0, \frac{T}{\lambda}]$ using the life cycle times R_n . The expectation $\tilde{\mathbb{E}}^{(\lambda)}[R_1 - R'_1]$ is smaller than $c > 0$ uniformly in $\lambda \leq 1$ by Lemma 4.6, so

$$\lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\int_0^{R_1} dr g_\lambda(S_r)\right] \leq \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\int_0^{R'_1} dr g_\lambda(S_r)\right] + c\lambda^{\frac{1}{2}}\left(\sup_{s \in S} g_\lambda(s)\right).$$

By Part (4) of Proposition 2.1 and since h has compact support, $g = \mathcal{A}_\lambda^+ + \mathbf{u}^{-1}h$ satisfies the conditions of Proposition 4.11. By Part (3) of Proposition 4.11, we have

$$\begin{aligned} \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\int_0^{R'_1} dr g_\lambda(S_r)\right] &= \lambda^{\frac{1}{2}} \int_{\tilde{\Sigma}} d\tilde{\mu}(x, p, z) \tilde{\mathbb{E}}^{(\lambda)}_{(x, p, z)}\left[\int_0^{R'_1} dr g_\lambda(S_r)\right] \\ &\leq C\lambda^{\frac{1}{2}} \int_{\tilde{\Sigma}} d\tilde{\mu}(x, p, z) (1 + \log(1 + |p|)) = C\lambda^{\frac{1}{2}} \int_{\Sigma} d\mu(x, p) (1 + \log(1 + |p|)) = O(\lambda^{\frac{1}{2}}), \end{aligned}$$

for some $C > 0$ and all $\lambda < 1$. The integral above is finite by our assumptions on the initial measure μ .

The other part of the first term on the right side of (4.56) is bounded by

$$\begin{aligned} \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\sup_{0 \leq n \leq \tilde{N}_{\frac{T}{\lambda}}} \int_{R_n}^{R_{n+1}} dr g_\lambda(S_r)\right] &\leq \lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}\left[\sum_{n=1}^{\tilde{N}_{\frac{T}{\lambda}}} \left(\int_{R_n}^{R_{n+1}} dr g_\lambda(S_r)\right)^4\right]^{\frac{1}{4}} \\ &\leq C^{\frac{1}{4}}\lambda^{\frac{1}{2}}\tilde{\mathbb{E}}^{(\lambda)}[\tilde{N}_{\frac{T}{\lambda}}]^{\frac{1}{4}} = O(\lambda^{\frac{3}{8}}). \end{aligned}$$

The first inequality uses that $\sup_n a_n \leq (\sum_n a_n^4)^{\frac{1}{4}}$ for positive numbers $a_n > 0$, and Jensen's inequality. The second inequality is for some $C > 0$ by the strong Markov property for the split chain $\tilde{\sigma}_n = S_{\tau_n}$ and by Part (1) of Proposition 4.11. The order equality is by our earlier result.

For the second term on the right side of (4.56), the key observation is that

$$\begin{aligned} b_\lambda &:= \tilde{\mathbb{E}}^{(\lambda)} \left[\int_{R_n}^{R_{n+1}} dr g'_\lambda(S_r) \mid \tilde{\mathcal{F}}_{R_{n-1}} \right] = \tilde{\mathbb{E}}^{(\lambda)} \left[\int_0^{R_1} dr g'_\lambda(S_r) \right] \\ &= \frac{\int_\Sigma ds \Psi_{\infty, \lambda}(s) (\mathcal{A}_\lambda^+(s) - \mathbf{u}^{-1}h(s))}{\int_\Sigma ds \Psi_{\infty, \lambda}(s) h(s)} = \frac{\int_\Sigma ds e^{-\lambda H(s)} (\mathcal{A}_\lambda^+(s) - \mathbf{u}^{-1}h(s))}{\int_\Sigma ds e^{-\lambda H(s)} h(s)} \end{aligned}$$

is $O(\lambda)$ for small λ . The first equality is by the strong Markov property for the chain $\tilde{\sigma}_n = \tilde{S}_{\tau_n}$ and by Part (1) of Proposition 4.1, and the second equality is by Part (2) of Proposition 4.4. The denominator of the rightmost expression approaches $\int_\Sigma ds h(s) = \mathbf{u}$, and the numerator is a difference of terms which are $1 + O(\lambda)$. This follows since $\int_\Sigma ds \mathbf{u}^{-1}h(s) = 1$, by the approximation $\int_\Sigma ds \mathcal{A}_\lambda^+(s) = 1 + O(\lambda)$ in Part (5) of Proposition 4.15, and since inserting the factor $e^{-\lambda H(s)}$ in the integrals will perturb the values by $O(\lambda)$.

With the above and the triangle inequality, the second term on the right side of (4.56) is smaller than

$$\begin{aligned} |b_\lambda| \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_{\frac{T}{\lambda}}] + \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left| \sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t + \frac{1}{2} \rfloor} \left(\int_{R_{2n-1}}^{R_{2n}} dr g'_\lambda(S_r) - b_\lambda \right) \right| \right] \\ + \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left| \sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t \rfloor} \left(\int_{R_{2n}}^{R_{2n+1}} dr g'_\lambda(S_r) - b_\lambda \right) \right| \right]. \end{aligned}$$

The first term is $O(\lambda)$, since $\lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_{\frac{T}{\lambda}}]$ is bounded for $\lambda \leq 1$. However,

$$\sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t + \frac{1}{2} \rfloor} \left(\int_{R_{2n-1}}^{R_{2n}} dr g'_\lambda(S_r) - b_\lambda \right) \quad \text{and} \quad \sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t \rfloor} \left(\int_{R_{2n}}^{R_{2n+1}} dr g'_\lambda(S_r) - b_\lambda \right)$$

are martingales with respect to the filtration $\tilde{\mathcal{F}}'_t$. Thus, we can apply the standard arguments to bound them:

$$\begin{aligned} \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left| \sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t + \frac{1}{2} \rfloor} \int_{R_{2n-1}}^{R_{2n}} dr g'_\lambda(S_r) - b_\lambda \right| \right] &\leq 2 \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sum_{n=1}^{\lfloor \frac{1}{2} \tilde{N}_t + \frac{1}{2} \rfloor} \left(\int_{R_{2n-1}}^{R_{2n}} dr g'_\lambda(S_r) - b_\lambda \right)^2 \right]^{\frac{1}{2}} \\ &\leq \lambda^{\frac{1}{2}} C^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_{\frac{T}{\lambda}}]^{\frac{1}{2}} = O(\lambda^{\frac{1}{4}}). \end{aligned}$$

The first inequality is Jensen's with the square function, and then Doob's maximal inequality. The second inequality is for some $C > 0$ by the strong Markov property applied at the time R_{2n-2} for the n th term in the sum followed by an application of Part (1) of Proposition 4.11.

Part (2): By Lemma 4.5, $\langle \tilde{M} \rangle_t$ is a sum of terms $\bar{v}_\lambda(S_{R_n})$, and so the difference between

$\langle \tilde{M} \rangle_t$ and $v_\lambda \tilde{N}_t$ can be written

$$\begin{aligned} \langle \tilde{M} \rangle_t - v_\lambda \tilde{N}_t &= \sum_{n=1}^{\tilde{N}_t} (\bar{v}_\lambda(S_{R_n}) - v_\lambda) = \bar{v}_\lambda(S_{R_1}) - \bar{v}_\lambda(S_{R_1}) + \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_{\tilde{N}_t}}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] + \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_1}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] \\ &\quad + \sum_{n=1}^{\tilde{N}_t} (\bar{v}_\lambda(S_{R_{n+1}}) - \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_n}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] + \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_{n+1}}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] - v_\lambda), \end{aligned}$$

where $\tilde{\delta}_s$ is the splitting of the δ -distribution at $s \in \Sigma$. Notice that $v_\lambda = \int_\Sigma d\nu(s) \bar{v}_\lambda(s)$. The sum on the right is martingale with respect to \tilde{F}'_t , by the same reasoning that \tilde{M}_t is a martingale. Since $S_{R_n} \in \text{Supp}(\nu)$ for $n \geq 1$, we have the standard inequalities:

$$\begin{aligned} \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left| \langle \tilde{M} \rangle_t - v_\lambda \tilde{N}_t \right| \right] &\leq 4\lambda^{\frac{1}{2}} \sup_{s \in \text{Supp}(\nu)} \bar{v}_\lambda(s) \\ &\quad + 2\lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\sum_{n=1}^{\tilde{N}_{\frac{T}{\lambda}}} \left(\bar{v}_\lambda(S_{R_{n+1}}) - \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_n}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] + \tilde{\mathbb{E}}_{\tilde{\delta}_{S_{R_{n+1}}}}^{(\lambda)} [\bar{v}_\lambda(S_{R_1})] - v_\lambda \right)^2 \right]^{\frac{1}{2}} \\ &\leq 4\lambda^{\frac{1}{2}} \sup_{s \in S} \bar{v}_\lambda(s) + 2^{\frac{3}{2}} \lambda^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} \left[\tilde{N}_{\frac{T}{\lambda}} \right]^{\frac{1}{2}} \left(\int_\Sigma ds \bar{v}_\lambda^2(s) - \left(\int_\Sigma ds \bar{v}_\lambda(s) \right)^2 \right)^{\frac{1}{2}} = O(\lambda^{\frac{1}{4}}). \end{aligned}$$

□

Theorem 4.18. *There exists a $C > 0$ such that for all $\lambda < 1$, then*

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} \left| \lambda^{\frac{1}{4}} D_{\frac{t}{\lambda}} \right| \right] \leq C.$$

In particular, there is convergence in probability

$$\sup_{0 \leq t \leq T} \left| \lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}} \right| \Longrightarrow 0.$$

Proof. We will work with the split dynamics to show that

$$\mathbb{E}^{(\lambda)} \left[\sup_{t \in [0, \frac{T}{\lambda}]} |D_t| \right] = \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{t \in [0, \frac{T}{\lambda}]} |D_t| \right] = O(\lambda^{-\frac{1}{4}}).$$

Let \tilde{M}_t be the martingale from Lemma 4.5. We can write D_t as

$$\begin{aligned} D_t &= \int_0^{R_1} dr \frac{dV}{dx}(X_r) - \int_t^{R_{\tilde{N}_t+1}} dr \frac{dV}{dx}(X_r) \\ &\quad + \tilde{\mathbb{E}}_{\tilde{S}_{R_1}} \left[\int_{R_1}^{R_2} dr \frac{dV}{dx}(X_r) \right] - \tilde{\mathbb{E}}_{\tilde{S}_{R_{\tilde{N}_t+1}}} \left[\int_{R_{\tilde{N}_t+1}}^{R_{\tilde{N}_t+2}} dr \frac{dV}{dx}(X_r) \right] + \tilde{M}_t. \end{aligned} \quad (4.57)$$

The triangle inequality gives

$$\begin{aligned} \mathbb{E}^{(\lambda)} \left[\sup_{t \in [0, \frac{T}{\lambda}]} |D_t| \right] &\leq 2\tilde{\mathbb{E}}^{(\lambda)} \left[\left| \int_0^{R_1} dr \frac{dV}{dx}(X_r) \right| \right] + 3\tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{R_1 \leq t \leq \frac{T}{\lambda}} \left| \int_t^{R_{\tilde{N}_t+1}} dr \frac{dV}{dx}(X_r) \right| \right] \\ &\quad + \mathbb{E}^{(\lambda)} \left[\sup_{t \in [0, \frac{T}{\lambda}]} |\tilde{M}_t| \right]. \end{aligned} \quad (4.58)$$

For the first term on the right side of (4.58), we can apply Part (2) of Proposition 4.14 to get

$$\begin{aligned} \tilde{\mathbb{E}}^{(\lambda)} \left[\left| \int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right| \right] &\leq \int d\tilde{\mu}(x, p, z) \tilde{\mathbb{E}}_{(x, p, z)}^{(\lambda)} \left[\left| \int_0^{R'_1} dr \frac{dV}{dx}(X_r) \right| \right] \\ &\leq C \int_{\mathbb{R}} d\mu(x, p) (1 + \log(1 + |p|)) < C \int_{\mathbb{R}} d\mu(x, p) (1 + |p|), \end{aligned} \quad (4.59)$$

where μ is the initial measure on Σ and $\tilde{\mu}$ is its splitting. By our assumption on μ , the first moment is finite, and hence (4.59) is finite.

The second term on the right side of (4.58) is less simple, since it depends on t . Writing

$$\int_t^{R_{\tilde{N}_{t+1}}} dr \frac{dV}{dx}(X_r) = \int_{R_{\tilde{N}_t}}^{R_{\tilde{N}_{t+1}}} dr \frac{dV}{dx}(X_r) - \int_{R_{\tilde{N}_t}}^t dr \frac{dV}{dx}(X_r),$$

then with the triangle inequality

$$\begin{aligned} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left| \int_t^{R_{\tilde{N}_{t+1}}} dr \frac{dV}{dx}(X_r) \right| \right] &\leq \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} \left(\left| \int_{R_{\tilde{N}_t}}^t dr \frac{dV}{dx}(X_r) \right| + \left| \int_{R_{\tilde{N}_t}}^{R_{\tilde{N}_{t+1}}} dr \frac{dV}{dx}(X_r) \right| \right) \right] \\ &\leq 2\tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{1 \leq m \leq \tilde{N}_t} \sup_{t \in [R_m, R_{m+1}]} \left| \int_{R_m}^t dr \frac{dV}{dx}(X_r) \right| \right]. \end{aligned} \quad (4.60)$$

We can get rid of the first supremum with the inequality

$$\begin{aligned} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{1 \leq m \leq \tilde{N}_{\frac{T}{\lambda}}} \sup_{t \in [R_m, R_{m+1}]} \left| \int_{R_m}^t dr \frac{dV}{dx}(X_r) \right| \right] &\leq \tilde{\mathbb{E}}^{(\lambda)} \left[\sum_{m=1}^{\tilde{N}_{\frac{T}{\lambda}}} \sup_{t \in [R_m, R_{m+1}]} \left| \int_{R_m}^t dr \frac{dV}{dx}(X_r) \right|^2 \right]^{\frac{1}{2}} \\ &\leq \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_{\frac{T}{\lambda}}]^{\frac{1}{2}} \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} \left[\sup_{t \in [0, R_1]} \left| \int_0^t dr \frac{dV}{dx}(X_r) \right|^2 \right]^{\frac{1}{2}} = O(\lambda^{-\frac{1}{4}}), \end{aligned} \quad (4.61)$$

where we have applied the strong Markov property at the times R_m and have used that \tilde{S}_{R_m} is distributed as $\tilde{\nu}$ by Part (1) of Proposition 4.1. By Part (1) of Proposition 4.14, the expectation $\tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}$ above is bounded uniformly for $\lambda < 1$. In the beginning of the proof of Lemma 4.17, we proved that $\tilde{\mathbb{E}}^{(\lambda)}[\tilde{N}_{\frac{T}{\lambda}}] = O(\lambda^{-\frac{1}{2}})$. Thus we have the last order equality above.

Now for the last term in (4.58). Applying Jensen's and then Doob's maximal inequality, we get the first two relations below

$$\begin{aligned} \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\tilde{M}_t| \right] &\leq \tilde{\mathbb{E}}^{(\lambda)} \left[\sup_{0 \leq t \leq \frac{T}{\lambda}} |\tilde{M}_t|^2 \right]^{\frac{1}{2}} \leq 2\tilde{\mathbb{E}}^{(\lambda)} [|\tilde{M}_{\frac{T}{\lambda}}|^2]^{\frac{1}{2}} \\ &= 2\tilde{\mathbb{E}}^{(\lambda)} [\langle \tilde{M} \rangle_{\frac{T}{\lambda}}]^{\frac{1}{2}} = 2v_{\lambda}^{\frac{1}{2}} \tilde{\mathbb{E}}^{(\lambda)} [\tilde{N}_{\frac{T}{\lambda}}]^{\frac{1}{2}}. \end{aligned} \quad (4.62)$$

The second equality is by Lemma 4.17 and $v_{\lambda} = \int_{\Sigma} d\nu(s) \overline{v}_{\lambda}(s)$. Since $\tilde{\mathbb{E}}^{(\lambda)}[\tilde{N}_{\frac{T}{\lambda}}] = O(\lambda^{-\frac{1}{2}})$ and v_{λ} is uniformly bounded for $\lambda < 1$, the proof is complete. \square

5 Proof of Theorem 1.2

Now we put together the results from previous sections to prove the convergence of $\lambda^{\frac{1}{2}} P_{\frac{t}{\lambda}} = P_t^{(\lambda)}$ to the Ornstein-Uhlenbeck process over the interval $[0, T]$.

Proof of Theorem 1.2. Theorem 4.18 contains the proof of the bound for the expectation of $\sup_{0 \leq t \leq \frac{T}{\lambda}} |D_t|$.

Let $P_t^{(\lambda),'}$ be defined as in Lemma 3.2. By Lemma 3.2, the difference $P_t^{(\lambda)} - P_t^{(\lambda),'}$ converges to zero with respect to the uniform metric over the interval $[0, T]$. Thus we can work with $P_t^{(\lambda),'}$ rather than $P_t^{(\lambda)}$. Define the map $\mathcal{G} : L^\infty([0, T]) \rightarrow L^\infty([0, T])$ given by

$$\mathcal{G}(h)_t = h_t - \frac{1}{2} \int_0^t dr e^{-\frac{1}{2}(t-r)} h_r, \quad h \in L^\infty([0, T]).$$

Notice that the solution \mathbf{p}_t to the Langevine equation (1.4) has the explicit solution

$$\mathbf{p}_t = \mathcal{G}(\mathbf{B}'_t), \quad (5.1)$$

where we have assumed $\mathbf{p}_0 = 0$. Moreover, the integral equation (3.1) for $P_t^{(\lambda),'}$ admits the closed form

$$P_t^{(\lambda),'} = e^{-\frac{1}{2}t} \lambda^{\frac{1}{2}} P_0 + \mathcal{G}(\lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}})_t + \mathcal{G}(M^{(\lambda)})_t. \quad (5.2)$$

By our assumption (2) of List 1.1, the moment $\mathbb{E}^{(\lambda)}[|P_0|]$ is finite, and thus the first term on the right side of (5.2) converges in probability to zero as $\lambda \rightarrow 0$. The random variable $\sup_{0 \leq t \leq T} |\mathcal{G}(\lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}})_t|$ converges in probability to zero also, since

$$\mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |\mathcal{G}(\lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}})_t| \right] \leq (1 + \frac{T}{2}) \mathbb{E}^{(\lambda)} \left[\sup_{0 \leq t \leq T} |\lambda^{\frac{1}{2}} D_{\frac{t}{\lambda}}| \right] = O(\lambda^{\frac{1}{4}}),$$

where the order equality follows by Theorem 4.18. By Lemma 3.3, $M_t^{(\lambda)}$ converges in law to a standard Brownian motion \mathbf{B}' with respect to the uniform metric. Since the map \mathcal{G} is continuous with respect to the supremum norm, $\mathcal{G}(M^{(\lambda)})_t$ converges in law to the process $\mathcal{G}(\mathbf{B}')_t$ with respect to the uniform metric. The process $\mathcal{G}(\mathbf{B}')_t$ is Ornstein-Uhlenbeck, and thus $P_t^{(\lambda),'}$ converges in law to the Ornstein-Uhlenbeck process and $P_t^{(\lambda)}$ does also. \square

6 Miscellaneous proofs

[Proof of Proposition 3.1.]

To ease the notations, we note $g(q) = (2\pi)^{-\frac{1}{2}} e^{-\frac{q^2}{2}}$. We have after the change of variables $q = \frac{\lambda+1}{2} p' + \frac{\lambda-1}{2} p = \frac{\lambda+1}{2} (p' - p) + \lambda p$

$$\mathcal{E}_\lambda(p) = \frac{2\eta}{\lambda+1} \int_{\mathbb{R}} dq |q - \lambda p| g(q) = \frac{2\eta}{\lambda+1} \left(2 \int_{\lambda p}^{\infty} q g(q) dq + \lambda p \int_{-\lambda p}^{+\lambda p} dq g(q) \right). \quad (6.1)$$

We have $\int_{\lambda p}^{\infty} qg(q) dq = g(\lambda p)$, and $\int_{-\lambda p}^{+\lambda p} dqg(q) \leq \min(1, \lambda|p|)$. By calculus, we also see that $\alpha \mapsto g(\alpha) + \alpha \int_0^{\alpha} dqg(q)$ has a minimum over \mathbb{R} at 0. Hence

$$\frac{1}{8(\lambda+1)} = \mathcal{E}_{\lambda}(0) \leq \mathcal{E}_{\lambda}(p) \leq \frac{4\eta g(0) + 2\eta \min(\lambda|p|, \lambda^2 p^2)}{(\lambda+1)} = \frac{(1 + C \min(\lambda|p|, \lambda^2|p|^2))}{8(\lambda+1)}. \quad (6.2)$$

We use exactly the same technique to estimate $\mathcal{D}_{\lambda}(p)$. Here we find

$$\mathcal{D}_{\lambda}(p) = \frac{4\eta}{(\lambda+1)^2} \left(- \int_{-\lambda p}^{+\lambda p} q^2 g(q) dq - 4\lambda p \int_{\lambda p}^{\infty} qg(q) dq - \lambda^2 p^2 \int_{-\lambda p}^{+\lambda p} g(q) dq \right) \quad (6.3)$$

$$= -\frac{2\lambda p}{\lambda+1} \mathcal{E}_{\lambda}(p) - \frac{4\eta}{(\lambda+1)^2} \int_{-\lambda p}^{+\lambda p} g(q) dq. \quad (6.4)$$

To prove the first inequality involving $\mathcal{D}_{\lambda}(p)$, we use (6.3). For the first integral, we note that for $|q| \leq \lambda|p|$, one has $q^2 g(q) \leq \lambda^2 p^2 g(q)$. We also observe that there exists a constant $C > 0$ such that $|g(q) - g(0)| \leq C|q|$ for all q , so

$$\mathcal{D}_{\lambda}(p) = -\frac{16\eta g(0)\lambda p}{(\lambda+1)^2} + O(\lambda^2 p^2).$$

The second inequality involving $\mathcal{D}_{\lambda}(p)$ is an obvious consequence of (6.4).

Now, we repeat the computation for $\Pi_{\lambda}^{(2)}(p)$ and we get

$$\begin{aligned} \Pi_{\lambda}^{(2)}(p) &= \frac{8\eta}{(\lambda+1)^3} \left((2\lambda^2 p^2 + 4)g(\lambda p) + \lambda p(3 + \lambda^2 p^2) \int_{-\lambda p}^{+\lambda p} g(q) dq \right) \\ &= \frac{4(\lambda^2 p^2 + 2)}{(\lambda+1)^2} \mathcal{E}_{\lambda}(p) + \frac{8\eta \lambda p}{(\lambda+1)^3} \int_{-\lambda p}^{+\lambda p} g(q) dq. \end{aligned} \quad (6.5)$$

Using (6.4), we have

$$\frac{4\lambda^2 p^2}{(\lambda+1)^2} \mathcal{E}_{\lambda}(p)^2 - \mathcal{D}_{\lambda}(p)^2 = -\frac{4\eta}{(\lambda+1)^2} \left(\int_{-\lambda p}^{+\lambda p} g(q) dq \right) \left(\frac{4\lambda p}{\lambda+1} \mathcal{E}_{\lambda}(p) + \frac{4\eta}{(\lambda+1)^2} \int_{-\lambda p}^{+\lambda p} g(q) dq \right).$$

Combining the two previous equality with (6.1), we find that

$$\mathcal{E}_{\lambda}(p) \mathcal{Q}_{\lambda}(p) = \frac{32\eta g(\lambda p)}{(\lambda+1)^3} \mathcal{E}_{\lambda}(p) + \frac{16\eta^2}{(\lambda+1)^4} \int_{-\lambda p}^{+\lambda p} g(q) dq \int_{-\lambda p}^{+\lambda p} (\lambda^2 p^2 - q^2) g(q) dq.$$

The first and the third equalities show that

$$\frac{32\eta g(\lambda p)}{(\lambda+1)^3} \leq \mathcal{Q}_{\lambda}(p) \leq \frac{8\mathcal{E}_{\lambda}(p)}{(\lambda+1)^2},$$

which is enough to show the upper bound for $\mathcal{Q}_{\lambda}(p)$ and the estimate for $\left| \mathcal{Q}_{\lambda}(p) - \frac{1}{(\lambda+1)^3} \right|$. For the lower bound, we deduce from the preceding calculations that $(\lambda+1)^3 \mathcal{Q}_{\lambda}(p)$ is actually a function of λp , which is continuous, strictly positive and goes to infinity as $\lambda p \rightarrow \pm\infty$, hence has a strictly positive minimum² not depending on λ .

²This minimum is not at 0.

Finally, reasoning as above, it is easy to produce an upper bound for $\Pi_\lambda^{(2m)}(p)$ which is a polynomial of degree $2m + 1$ in $\lambda|p|$.

The inequalities in Part 7 concern the moments for the momentum distribution when conditioned to jump from a momentum p in one region to a momentum p' in another region. Both inequalities follow from the monotonic and sharply decaying form of the tails of the rates $r_{\lambda,p}(p') = \mathcal{J}_\lambda(p, p')$.

[Proof of Proposition 4.3.]

We will show (1). Let $(\mathbf{x}_t(x, p), \mathbf{p}_t(x, p))$ be the trajectory starting from the phase-space point (x, p) and evolving according to the Hamiltonian $H(x, p) = \frac{1}{2}p^2 + V(x)$. The kernel \mathcal{T}_λ satisfies the following closed integral equation

$$\begin{aligned} \mathcal{T}_\lambda(x, p; dx', dp') &= \int_0^\infty dt \delta(\mathbf{x}_t(x, p) - x', \mathbf{p}_t(x, p) - p') e^{-\int_0^t ds (1 + \mathcal{E}_\lambda(\mathbf{p}_s(x, p)))} \\ &\quad + \int_0^\infty dt \int_{\mathbb{R}} dp'' \frac{\mathcal{J}_\lambda(\mathbf{p}_t(x, p), p'')}{\mathcal{E}_\lambda(\mathbf{p}_t(x, p))} \mathcal{T}_\lambda(\mathbf{x}_t(x, p), p''; dx', dp') e^{-\int_0^t ds (1 + \mathcal{E}_\lambda(\mathbf{p}_s(x, p)))}. \end{aligned}$$

By iterating this equation, we obtain a series expansion for \mathcal{T}_λ where the n th term corresponds to the event that $n - 1$ collisions occur over the time interval $[0, t]$. By considering the contribution to the transition kernel \mathcal{T}_λ due to a single collision occurring over the mean-1 exponential time interval, we have the following lower bound

$$\begin{aligned} &\mathcal{T}_\lambda(x, p; dx', dp') \\ &\geq dx' dp' \int_{\mathbb{R}} dp'' \int_0^\infty dt \int_0^{t_2} dt_2 \int_0^{t_2} dt_1 \frac{\mathcal{J}_\lambda(\mathbf{p}_{t_1}(x, p), p'')}{\mathcal{E}_\lambda(\mathbf{p}_{t_1}(x, p))} \frac{\mathcal{J}_\lambda(\mathbf{p}_{t_2-t_1}(\mathbf{x}_{t_1}(x, p), p''), \mathbf{p}_{-(t-t_2)}(x', p'))}{\mathcal{E}_\lambda(\mathbf{p}_{t_2-t_1}(\mathbf{x}_{t_1}(x, p), p''))} \\ &\quad e^{-\int_0^{t-t_2} ds (1 + \mathcal{E}_\lambda(\mathbf{p}_{-s}(x', p')))} e^{-\int_0^{t_2-t_1} ds (1 + \mathcal{E}_\lambda(\mathbf{p}_s(\mathbf{x}_{t_1}(x, p), p'')))} e^{-\int_0^{t_1} ds (1 + \mathcal{E}_\lambda(\mathbf{p}_s(x, p)))}. \quad (6.6) \end{aligned}$$

Let $A \subset \mathbb{R}$ be the set of p with $2l \leq \frac{1}{2}p^2 \leq 5l$. Let $c > 0$ be the minimum of the values

$$\inf_{\lambda < 1} \inf_{p^2 \leq l, p' \in A} \frac{\mathcal{J}_\lambda(p, p')}{\mathcal{E}_\lambda(p)} \quad \text{and} \quad \inf_{\lambda < 1} \inf_{p^2 \leq l, p' \in A} \frac{\mathcal{J}_\lambda(p', p)}{\mathcal{E}_\lambda(p')}.$$

We have defined the set A to exclude the arguments p, p' of \mathcal{J}_λ from being close, since the rates $\mathcal{J}_\lambda(p, p')$ have a zero along the line $p' = p$. If $H(x, p), H(x', p') \leq l$, the conservation of energy guarantees that $\frac{1}{2}\mathbf{p}_{t_1}^2(x, p) \leq l$ and $\frac{1}{2}\mathbf{p}_{-(t-t_2)}^2(x', p') \leq l$. Also, if $3l \leq \frac{1}{2}(p'')^2 \leq 4l$, then $\mathbf{p}_{t_2-t_1}(\mathbf{x}_{t_1}(x, p), p'') \in A$, since the kinetic energy can not fluctuate by more than l through the Hamiltonian evolution. For all $(x, p), (x', p')$ with $H(x, p), H(x', p') \leq l$ and all $\lambda < 1$,

$$\mathcal{T}_\lambda(x, p; dx', dp') \geq c^2 dx' dp' \left(\int_A dp'' \right).$$

Thus, we can take $\mathbf{u} = c^2 U^2 \left(\int_A dp'' \right)$.

Now, we prove the upper bound for $\mathcal{T}_\lambda(s, ds')$. Notice that for $\Psi \in L^1(\Sigma) \cap L^\infty(\Sigma)$, then $\|\mathcal{T}_\lambda(\Psi)\|_\infty \leq \|\Psi\|_\infty$. In other words, \mathcal{T}_λ is a contraction in the supremum norm. This is evident

from the resolvent form $\mathcal{T}_\lambda = \int_0^\infty dt e^{-t} \Phi_{t,\lambda}$, where $\Phi_{t,\lambda}$ are the dynamical maps for the Master equation (1.5) and by the inequalities

$$\|\mathcal{T}_\lambda \Psi\|_\infty \leq \int_0^\infty dt e^{-t} \|\Phi_{t,\lambda}(\Psi)\|_\infty \leq \|\Psi\|_\infty.$$

The maps $\Phi_{t,\lambda}$ are contractive in the supremum norm, since the dynamics is driven by Hamiltonian flow which preserves the supremum norm, and noise satisfying the detailed balance condition

$$e^{-\frac{1}{2}p_1^2} \mathcal{J}_\lambda(p_1, p_2) = e^{-\frac{1}{2}p_2^2} \mathcal{J}_\lambda(p_2, p_1).$$

When $H(s') \neq H(s)$, then a collision must occur over the time interval $[0, \tau_1]$ in order for $S_0 = s$ and $S_{\tau_1} = s'$. Considering the event that the first collision occurs before time τ_1 , then the strong Markov property gives the first equality below:

$$\mathcal{T}_\lambda(s, ds') = \mathbb{E}_s^{(\lambda)} \left[\chi(t_1 \leq \tau_1) \mathcal{T}_\lambda(S_{t_1}, ds') \right] \leq \|D_{\lambda,s}\|_\infty, \quad (6.7)$$

where $D_s^{(\lambda)}$ is the probability density of the first collision when starting from $s \in \Sigma$. The density has the closed form

$$D_s^{(\lambda)}(s') = \int_0^\infty dt \delta(\mathbf{x}_t(s) - x') \mathcal{J}_\lambda(\mathbf{p}_t(s), p') e^{-\int_0^t dr \mathcal{E}_\lambda(\mathbf{p}_r(s))}.$$

When $H(s) \geq 1 + 2 \sup_x V(x)$, the particle will revolve around the torus freely with speed $|p| \geq (2 + 2 \sup_x V(x))^{\frac{1}{2}}$. Using the above form for $D_s^{(\lambda)}$:

$$\begin{aligned} D_s^{(\lambda)}(s') &\leq \left(\sup_{p,p'} \frac{\mathcal{J}_\lambda(p, p')}{\mathcal{E}_\lambda(p)} \right) \int_0^\infty dt \delta(\mathbf{x}_t(s) - x') e^{-\int_0^t dr \mathcal{E}_\lambda(\mathbf{p}_r(s))} \mathcal{E}_\lambda(\mathbf{p}_t(s)) \\ &\leq \left(\sup_{p,p'} \frac{\mathcal{J}_\lambda(p, p')}{\mathcal{E}_\lambda(p)} \right) \left(\sum_{n=1}^\infty \frac{\mathcal{E}_\lambda(q(s, x'))}{q(s, x')} e^{-n \int_{\mathbb{T}} da \frac{\mathcal{E}_\lambda(q(s, a))}{q(s, a)}} \right) \\ &\leq \left(\sup_{p,p'} \frac{\mathcal{J}_\lambda(p, p')}{\mathcal{E}_\lambda(p)} \right) \left(\frac{\frac{\mathcal{E}_\lambda(q(s, x'))}{q(s, x')}}{1 - e^{-\int_{\mathbb{T}} da \frac{\mathcal{E}_\lambda(q(s, a))}{q(s, a)}}} \right), \end{aligned}$$

where $q(s, a) = 2^{\frac{1}{2}}(H(s) - V(a))^{\frac{1}{2}}$. The two terms on the right are bounded uniformly for all $\lambda < 1$ and s with $H(s) \geq l$.

[Proof of Propositions 2.1 and 4.15.]

Our first observation is that $\mathcal{A}_\lambda(x, p) = \mathcal{A}_\lambda(x, -p)$, $\mathcal{E}_\lambda(p) = \mathcal{E}_\lambda(-p)$ and $\mathcal{V}_\lambda(x, p) = \mathcal{V}_\lambda(x, -p)$. So for all the inequalities we have to prove, we can assume without loss of generality $p > 0$.

Assume now that $\lambda^{-\frac{3}{8}} \leq p \leq \lambda^{-\frac{3}{4}}$. We note that

$$H(x, p')^{\frac{1}{2}} - H(x, p)^{\frac{1}{2}} = \frac{\frac{p'^2}{2} - \frac{p^2}{2}}{\left(\frac{p'^2}{2} + V(x) \right)^{\frac{1}{2}} + \left(\frac{p^2}{2} + V(x) \right)^{\frac{1}{2}}} = \frac{1}{\sqrt{2}}(p' - p)\Gamma(p', p),$$

where $\Gamma(p', p) := 2^{-\frac{1}{2}}(p + p') \left(\left(\frac{p'^2}{2} + V(x) \right)^{\frac{1}{2}} + \left(\frac{p^2}{2} + V(x) \right)^{\frac{1}{2}} \right)^{-1}$. Thus one can write

$$\mathcal{A}_\lambda(x, p) = \mathcal{D}_\lambda(p) - \int_{\mathbb{R}} dp' (p' - p)(1 - \Gamma(p, p')) \mathcal{J}_\lambda(p, p'). \quad (6.8)$$

We know from Proposition 3.1 that

$$\left| \mathcal{D}_\lambda(p) + \frac{\lambda p}{2} \right| \leq C_1 \lambda^2 p^2 + C_2 \lambda^2 |p| \leq C \lambda^{\frac{5}{4}} |p|.$$

To estimate the integral in (6.8), we split it into two parts: $|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}$ and $|p' - \tilde{p}| < \frac{\tilde{p}}{2}$ where $\tilde{p} = \frac{1-\lambda}{1+\lambda}p$ is the center of the Gaussian density in \mathcal{J}_λ . The first part is actually in the Gaussian tail and goes to 0 at exponential speed when $\lambda \rightarrow 0$. Indeed, the quantity $|\Gamma(p, p')|$ is bounded by 1 and we have after the change of variables $q = \frac{\lambda+1}{2}p' + \frac{\lambda-1}{2}p$,

$$\int_{|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}} dp' |p' - p| \mathcal{J}_\lambda(p, p') = \frac{4\eta}{(\lambda+1)^2} \int_{|q| \geq \frac{(1-\lambda)}{4}p} dq (q - \lambda p)^2 \frac{e^{-\frac{q^2}{2}}}{\sqrt{2\pi}} \quad (6.9)$$

Since $\frac{(1-\lambda)}{4}p \rightarrow \infty$ as $\lambda \rightarrow 0$, we can bound this by, say, $C' \lambda^{\frac{5}{4}} |p|$. The remaining part $|p - \tilde{p}| < \frac{\tilde{p}}{2}$ of the integral in (6.8) can be as large as λp if p is not large enough for fixed λ . Now we show that this does not happen if $p \geq \lambda^{-\frac{3}{8}}$. First, we note that the function $\Gamma(p, p')$ is symmetric and it is non-decreasing in p' . More precisely, for $\frac{\tilde{p}}{2} \leq p' \leq \frac{3\tilde{p}}{2}$ and λ small enough,

$$0 \leq \frac{\partial \Gamma(p, p')}{\partial p'} = \frac{(p^2 + 2V(x))^{\frac{1}{2}}(p'^2 + 2V(x))^{\frac{1}{2}} - pp' + 2V(x)}{(p'^2 + 2V(x))^{\frac{1}{2}} \left((p^2 + 2V(x))^{\frac{1}{2}} + (p'^2 + 2V(x))^{\frac{1}{2}} \right)^2} \leq \frac{C}{p^3}. \quad (6.10)$$

We also have for $\frac{\tilde{p}}{2} \leq p'$ and λ small enough

$$0 \leq 1 - \Gamma(p, p') \leq 1 - \Gamma\left(\frac{p}{4}, \frac{p}{4}\right) = 1 - \frac{p}{(p^2 + 32V(x))^{\frac{1}{2}}} \leq \frac{C}{p^2}. \quad (6.11)$$

Using (6.11) the fact $p - \tilde{p} = \frac{2\lambda p}{\lambda+1}$ and (6.10), we can write (with $r = p' - \tilde{p}$)

$$\begin{aligned} & \int_{\frac{\tilde{p}}{2}}^{\frac{3\tilde{p}}{2}} dp' (p' - p)(1 - \Gamma(p, p')) \mathcal{J}_\lambda(p, p') \\ &= \frac{\eta(1+\lambda)}{2} \int_{-\frac{\tilde{p}}{2}}^{+\frac{\tilde{p}}{2}} dr (\tilde{p} - p + r) |\tilde{p} - p + r| (1 - \Gamma(p, \tilde{p} + r)) \frac{\exp\left(-\frac{1}{2} \left(\frac{(\lambda+1)}{2}r\right)^2\right)}{\sqrt{2\pi}} \\ &= -\frac{\eta(1+\lambda)}{2} \int_0^{+\frac{\tilde{p}}{2}} dr r^2 (\Gamma(p, \tilde{p} + r) - \Gamma(p, \tilde{p} - r)) \frac{\exp\left(-\frac{1}{2} \left(\frac{(\lambda+1)}{2}r\right)^2\right)}{\sqrt{2\pi}} + O(\lambda^{\frac{7}{4}}p) \\ &= O\left(\frac{1}{p^3}\right) + O(\lambda^{\frac{7}{4}}p). \end{aligned}$$

For $p \geq \lambda^{-\frac{3}{8}}$, the above can be bounded by $C \lambda^{\frac{3}{2}} p$. Note by the way that $\Gamma(p, \tilde{p} + r) \geq \Gamma(p, \tilde{p} - r)$, so the corresponding integral is strictly positive. Since $\frac{\partial \Gamma}{\partial p'}(p, p) = V(x)(p^2 + 2V(x))^{-\frac{3}{2}} \sim C p^{-3}$,

this means that the contribution of the integral in (6.8) can be of the same order as $\mathcal{D}_\lambda(p)$ if p is not large enough for fixed λ . Finally, collecting the above estimate, we have proved that for $\lambda^{-\frac{3}{8}} \leq p \leq \lambda^{-\frac{1}{4}}$, and λ small enough,

$$\left| \mathcal{A}_\lambda(x, p) + \frac{\lambda p}{2} \right| \leq C \lambda^{\frac{5}{4}} p.$$

This is enough to show that in this case, $\mathcal{A}_\lambda(x, p) < 0$ for λ small enough, so the proof of our first inequality is done.

Our previous analysis implies that for $\lambda^{-\frac{3}{8}} \leq p \leq \lambda^{-\frac{1}{4}}$,

$$\frac{(\mathcal{A}_\lambda(x, p))^2}{\mathcal{E}_\lambda(p)} = \frac{2\lambda^2 p^2}{\lambda + 1} + O(\lambda^{\frac{9}{4}} p^2) = O(\lambda^2 p^2) = O(\lambda^{\frac{1}{2}}).$$

Using the definition of Γ , we have

$$\begin{aligned} \mathcal{V}_\lambda(x, p) &= 2\mathcal{K}_{\lambda,2}(x, p) - \frac{(\mathcal{A}_\lambda(x, p))^2}{\mathcal{E}_\lambda(p)} \\ &= \Pi_\lambda^{(2)}(p) - \frac{(\mathcal{A}_\lambda(x, p))^2}{\mathcal{E}_\lambda(p)} + \int_{\mathbb{R}} dp' (p' - p)^2 (\Gamma(p, p')^2 - 1) \mathcal{J}_\lambda(p, p'). \end{aligned} \quad (6.12)$$

From (6.5), we see that

$$\Pi_\lambda^{(2)}(p) = 1 + O(\lambda^{\frac{1}{2}}).$$

The analysis of the remaining integral in (6.12) is done as before. More precisely, when $|p' - \tilde{p}| \leq \frac{\tilde{p}}{2}$, we can still apply (6.11) and we have

$$0 \leq 1 - \Gamma(p, p')^2 = (1 - \Gamma(p, p'))(1 + \Gamma(p, p')) \leq \frac{2C}{p^2}.$$

Next, we write

$$\int_{\frac{\tilde{p}}{2}}^{\frac{3\tilde{p}}{2}} dp' (p' - p)^2 (1 - \Gamma(p, p')^2) \mathcal{J}_\lambda(p, p') \leq \frac{8\eta}{(\lambda + 1)^3} \int_{|q| \leq \frac{(1-\lambda)}{4} p} |q - \lambda p|^3 \frac{2C}{p^2} \frac{e^{-\frac{q^2}{2}}}{\sqrt{2\pi}}. \quad (6.13)$$

Since $|q - \lambda p|^3 \leq (|q| + 1)^3$ and $p^{-2} \leq \lambda^{\frac{3}{4}}$, the above integral is $O(\lambda^{\frac{1}{2}})$. Finally, the integral in (6.12) over $|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}$ goes to 0 when $\lambda \rightarrow 0$ at exponential speed in λ , so is also $O(\lambda^{\frac{1}{2}})$. Collecting the previous estimates, we find that $|\mathcal{V}_\lambda(x, p) - 1| \leq C\lambda^{\frac{1}{2}}$ and $|2\mathcal{K}_{\lambda,2}(x, p) - 1| \leq C\lambda^{\frac{1}{2}}$ for λ small enough and $p^{-\frac{3}{8}} \leq p \leq \lambda^{-\frac{1}{4}}$.

For the proof of $\mathcal{A}_\lambda^-(x, p) \leq |\mathcal{D}_\lambda(p)|$, we consider $p > 0$. Then we have to show that the integral in (6.8) is nonpositive. For this, we use the monotonicity of $\Gamma(p, p')$ and the following straightforward inequality (valid for all $p, r \geq 0$),

$$\mathcal{J}_\lambda(p, p + r) \leq \mathcal{J}_\lambda(p, p - r).$$

Then we have

$$\begin{aligned} &\int_{\mathbb{R}} dp' (p' - p) (1 - \Gamma(p, p')) \mathcal{J}_\lambda(p, p') \\ &= \int_0^\infty r dr (1 - \Gamma(p, p + r)) \mathcal{J}_\lambda(p, p + r) + \int_0^\infty r dr (\Gamma(p, p - r) - 1) \mathcal{J}_\lambda(p, p - r) \\ &\leq \int_0^\infty r dr (\Gamma(p, p - r) - \Gamma(p, p + r)) \mathcal{J}_\lambda(p, p - r) \leq 0. \end{aligned}$$

In order to prove that $\frac{\mathcal{A}_\lambda(x,p)}{\mathcal{E}_\lambda(p)} + \frac{2\lambda|p|}{1+\lambda}$ is bounded, it is again enough to consider the case $p > 0$ since $\mathcal{E}_\lambda(p) = \mathcal{E}_\lambda(-p)$. Moreover, in view of Proposition 3.1 and (6.8), all we have to do is to prove that the integral in (6.8) is bounded by $C + C'\lambda p$ for suitable C, C' (actually, our analysis shows that one can take $C' = 0$). If $p \leq \lambda^{-1}$, then $(q - \lambda p)^2 \leq (|q| + 1)^2$. So we can do a change of variable as in (6.9) and we get that the integral in (6.8) is bounded in this case. Now, we assume that $p \geq \lambda^{-1}$. For further use, we prove a stronger result, namely that in this case, the integral in (6.8) goes to 0 as $\lambda \rightarrow 0$. Indeed, we can split this integral as before into $|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}$ and $|p' - \tilde{p}| < \frac{\tilde{p}}{2}$. The $|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}$ part is again in the Gaussian tail hence goes to 0 as $\lambda \rightarrow 0$ (see (6.9)). Finally, for the $|p' - \tilde{p}| \leq \frac{\tilde{p}}{2}$ part, we use (6.11) and we get

$$\int_{\frac{\tilde{p}}{2}}^{\frac{3\tilde{p}}{2}} dp' |(p' - p)(1 - \Gamma(p, p'))| \mathcal{J}_\lambda(p, p') \leq \frac{4\eta}{(\lambda + 1)^2} \int_{|q| \leq \frac{(1-\lambda)p}{4}} dq \frac{C}{p^2} (q - \lambda p)^2 \frac{e^{-\frac{q^2}{2}}}{\sqrt{2\pi}}.$$

Since $p \geq \lambda^{-1}$, we have $p^{-2}(q - \lambda p)^2 \leq \lambda^2(|q| + 1)^2$, and the above integral goes to 0 as $\lambda \rightarrow 0$. So the proof is complete.

Now we want to find a lower bound for $\frac{\mathcal{A}_\lambda^-(x,p)}{\mathcal{E}_\lambda(p)}$ for $|p| \geq \lambda^{-1}$. It is still enough to consider $p > 0$. Then by (6.8) and (6.4), we have for all $p > 0$

$$\frac{\mathcal{A}_\lambda(x,p)}{\mathcal{E}_\lambda(p)} \leq -\frac{2\lambda p}{\lambda + 1} - \int_{\mathbb{R}} dp' (p' - p)(1 - \Gamma(p, p')) \frac{\mathcal{J}_\lambda(p, p')}{\mathcal{E}_\lambda(p)}.$$

For $p \geq \lambda^{-1}$, we have $-\frac{2\lambda p}{\lambda + 1} \leq -1$. Since we have just proved that the remaining integral goes to 0 as $\lambda \rightarrow 0$ and $p \geq \lambda^{-1}$, we conclude that for λ small enough, $\frac{\mathcal{A}_\lambda(x,p)}{\mathcal{E}_\lambda(p)} \leq -\frac{1}{2}$.

In Proposition 3.1, we proved that $\mathcal{Q}_\lambda(p) \leq C\mathcal{E}_\lambda(p)$. Thus, to prove the same thing for $\mathcal{V}_\lambda(x,p)$, we have to prove that there are constants C, C' such that for all $p > 0$, both $\frac{\mathcal{D}_\lambda^2(p)}{\mathcal{E}_\lambda(p)} - \frac{\mathcal{A}_\lambda^2(x,p)}{\mathcal{E}_\lambda(p)}$ and the remaining integral in (6.12) are bounded by $C + C'\lambda p$. First, we write $\mathcal{D}_\lambda^2 - \mathcal{A}_\lambda^2 = (\mathcal{D}_\lambda - \mathcal{A}_\lambda)(\mathcal{D}_\lambda + \mathcal{A}_\lambda)$. We already proved (in the proof that $\frac{\mathcal{A}_\lambda(x,p)}{\mathcal{E}_\lambda(p)} + \frac{2\lambda|p|}{1+\lambda}$ is bounded) that $\mathcal{D}_\lambda(p) - \mathcal{A}_\lambda(x,p)$ (which is the integral in (6.8)) is bounded. We also proved above and in Proposition 3.1 that both $\frac{\mathcal{D}_\lambda(p)}{\mathcal{E}_\lambda(p)}$ and $\frac{\mathcal{A}_\lambda(x,p)}{\mathcal{E}_\lambda(p)}$ are bounded by $\frac{2\lambda|p|}{1+\lambda} + C'$ for some C' . Regarding the integral in (6.12), we consider two cases: $p \leq \lambda^{-1}$ and $p \geq \lambda^{-1}$. We have

$$\int_{\mathbb{R}} dp' (p' - p)^2 (1 - \Gamma(p, p')^2) \mathcal{J}_\lambda(p, p') \leq \frac{8\eta}{(\lambda + 1)^3} \int_{\mathbb{R}} |q - \lambda p|^3 \frac{e^{-\frac{q^2}{2}}}{\sqrt{2\pi}}.$$

If $p \leq \lambda^{-1}$, then $|q - \lambda p|^3 \leq (|q| + 1)^3$, so the integral is bounded. If $p \geq \lambda^{-1}$, then the integral over $|p' - \tilde{p}| \geq \frac{\tilde{p}}{2}$ decays exponentially when $\lambda \rightarrow 0$, so *a fortiori* is bounded. Finally, when $p \geq \lambda^{-1}$, for the integral over $|p' - \tilde{p}| < \frac{\tilde{p}}{2}$, we use (6.13). In this case, we have $|q - \lambda p|^3 p^{-2} \leq \lambda^2 |q - \lambda p| (|q| + 1)^2$. So we can bound the integral by $C + C'\lambda p$.

Now we prove the upper bound for $\mathcal{A}_\lambda^+(x,p)$. We assume $p \geq 1$. Recall that $p - \tilde{p} = \frac{2\lambda p}{\lambda + 1}$ so we can write

$$\mathcal{A}_\lambda(x,p) = \int_{\mathbb{R}} dp' (p' - \tilde{p}) \Gamma(p, p') \mathcal{J}_\lambda(p, p') - \frac{2\lambda p}{\lambda + 1} \int_{\mathbb{R}} dp' \Gamma(p, p') \mathcal{J}_\lambda(p, p').$$

But we have

$$-\frac{2\lambda p}{\lambda + 1} \int_{\mathbb{R}} dp' \Gamma(p, p') \mathcal{J}_\lambda(p, p') \leq -\frac{2\lambda p}{\lambda + 1} \int_{-\infty}^{-p} dp' \Gamma(p, p') \mathcal{J}_\lambda(p, p') \leq \frac{2\lambda p}{\lambda + 1} \int_{-\infty}^{-p} dp' \mathcal{J}_\lambda(p, p').$$

This is exponentially decreasing in p (uniformly in λ). Now, we write ($r = p' - \tilde{p}$ and $g(q) = (2\pi)^{-\frac{1}{2}} e^{-\frac{q^2}{2}}$)

$$\begin{aligned}
& \int_{\mathbb{R}} dp' (p' - \tilde{p}) \Gamma(p, p') \mathcal{J}_\lambda(p, p') \frac{2}{\eta(\lambda + 1)} \\
&= \int_0^\infty r dr \left(\Gamma(p, \tilde{p} + r) \left| r - \frac{2\lambda p}{\lambda + 1} \right| - \Gamma(p, \tilde{p} - r) \left| r + \frac{2\lambda p}{\lambda + 1} \right| \right) g\left(\frac{\lambda + 1}{2} r\right) \\
&\leq \int_0^{\frac{2\lambda p}{\lambda + 1}} r dr \frac{2\lambda p}{\lambda + 1} (\Gamma(p, \tilde{p} + r) - \Gamma(p, \tilde{p} - r)) g\left(\frac{\lambda + 1}{2} r\right) \\
&+ \int_{\frac{2\lambda p}{\lambda + 1}}^\infty r^2 dr (\Gamma(p, \tilde{p} + r) - \Gamma(p, \tilde{p} - r)) g\left(\frac{\lambda + 1}{2} r\right).
\end{aligned}$$

In this last inequality, we used the fact that $\Gamma(p, \tilde{p} + r) + \Gamma(p, \tilde{p} - r) \geq 0$, which is true because $\Gamma(p, \tilde{p} - r) \geq \Gamma(p, -\tilde{p} - r)$ and $|\Gamma(p, -\tilde{p} - r)| \leq \Gamma(p, \tilde{p} + r)$ for all $r \geq 0$. Now, for the first integral, we can use (6.10) and we find that it is bounded by $\frac{C\lambda}{p^2}$. For the second integral, we first observe that the integral over $r \geq \frac{\tilde{p}}{2}$ is decaying exponentially, and for the integral over $0 \leq r \leq \frac{\tilde{p}}{2}$, we use again (6.10) and we get a $\frac{C}{p^3}$ bound.

The bounds for $\mathcal{K}_{\lambda,n}$ are straightforward since by definition of Γ , one has

$$\mathcal{K}_{\lambda,n}(x, p) \leq 2^{-\frac{n}{2}} \int_{\mathbb{R}} dp' |p - p'|^n \mathcal{J}_\lambda(p, p').$$

In the same way, we have

$$\mathcal{K}_{\lambda,n}(x, p) \leq 2^{-\frac{n}{2}} \int_{\mathbb{R}} dp' |p - p'|^n \chi(|p'| > |p|) \mathcal{J}_\lambda(p, p').$$

Since $\mathcal{K}_{\lambda,n}(x, -p) = \mathcal{K}_{\lambda,n}(x, p)$, we may assume without loss of generality that $p > 0$. Writing $r = p' - p$, we have

$$\begin{aligned}
\mathcal{K}_{\lambda,n}(x, p) &\leq C_n \left(\int_0^\infty r^{m+1} dr g\left(\frac{\lambda + 1}{2} r + \lambda p\right) + \int_{2p}^\infty r^{m+1} dr g\left(\frac{\lambda + 1}{2} r - \lambda p\right) \right) \\
&\leq C_n \left(\int_0^\infty r^{m+1} dr g\left(\frac{\lambda + 1}{2} r\right) + \int_{2p}^\infty r^{m+1} dr g\left(\frac{r}{2}\right) \right).
\end{aligned}$$

Therefore, $\mathcal{K}_{\lambda,n}(x, p)$ is bounded by some $C_n > 0$.

Acknowledgments

The authors are grateful to Professor Höpfner for sending a copy of Touati's unpublished paper and giving helpful comments. This work is supported by the European Research Council grant No. 227772.

A Exponential ergodicity

In this section, we present for convenience a self-contained proof of the relaxation of our dynamics to an equilibrium state. Our dynamics is driven by the forward Kolmogorov equation

$$\begin{aligned} \frac{d}{dt}\Psi_{t,\lambda}(x, p) &= (\mathcal{L}_\lambda^* \Psi_{t,\lambda})(x, p) = -p \frac{\partial}{\partial x} \Psi_{t,\lambda}(x, p) + \frac{dV}{dx}(x) \frac{\partial}{\partial p} \Psi_{t,\lambda}(x, p) \\ &\quad + \int_{\mathbb{R}} dp' (\mathcal{J}_\lambda(p', p) \Psi_{t,\lambda}(x, p') - \mathcal{J}_\lambda(p, p') \Psi_{t,\lambda}(x, p)), \end{aligned} \quad (\text{A.1})$$

which has equilibrium state $\Psi_{\infty,\lambda}$. The Kolmogorov equation (A.1) determines a transition semigroup $\Phi_{t,\lambda}$, which we take to operate on measures from the right and bounded measurable functions from the left:

$$(\mu \Phi_{t,\lambda})(ds) = \int_{\Sigma} \mu(ds') \Phi_{t,\lambda}(s', ds) \quad \text{and} \quad (\Phi_{t,\lambda} g)(s) = \int_{\Sigma} \Phi_{t,\lambda}(s, ds') g(s'),$$

for $\mu \in M(\Sigma)$ and $g \in B(\Sigma)$. We identify probability densities with their corresponding measures and denote the total variation norm for finite measures on Σ by $\|\cdot\|_1$. The exponential ergodicity that we prove in Theorem A.1 is not used critically anywhere in our proofs, although we need some degree of ergodicity in order to make sense of certain expressions such as the reduced resolvent \mathfrak{R}_λ of the function $\frac{dV}{dx}$, for instance. A more thorough study of the ergodicity would give some control of the exponential rate $\alpha(\lambda)$ appearing in Theorem A.1 for $\lambda \ll 1$. We believe that $\alpha(\lambda)$ can be taken to scale as $\propto \frac{\lambda}{\log(\lambda^{-1})}$ for small λ (i.e. a bit slower than linear in λ), and if C is replaced on the right side of (A.2) by $C\|\Psi\|_w$ for the weighted norm $\|\Psi\|_w = \int_{\Sigma} |\Psi|(dx dp) (1 + \lambda^{\frac{1}{2}}|p|)$, then the exponential rate scales as $\propto \lambda$ for $\lambda \ll 1$.

We denote the space of probability measures on Σ by $M_+^1(\Sigma)$.

Theorem A.1. *Let $\Phi_{t,\lambda}$ be the transition semigroup corresponding to the Kolmogorov equation (A.1). There exist $\alpha(\lambda), C > 0$ such that for all $\Psi \in M_+^1(\Sigma)$,*

$$\|\Psi \Phi_{t,\lambda} - \Psi_{\infty,\lambda}\|_1 \leq C e^{-t\alpha(\lambda)}. \quad (\text{A.2})$$

Proof. It is sufficient to work with the resolvent chain rather than the original process and show that there are $C, \alpha > 0$

$$\begin{aligned} C e^{-n\alpha} &\geq \|\Psi \mathcal{T}_\lambda^n - \Psi_{\infty,\lambda}\|_1 \\ &= \sup_{\|g\|_\infty \leq 1} \left| \int_{\Sigma} (\Psi \mathcal{T}_\lambda^n)(ds) g(s) - \Psi_{\infty,\lambda}(g) \right|, \end{aligned} \quad (\text{A.3})$$

where $\mathcal{T}_\lambda : B(\Sigma) \rightarrow B(\Sigma)$ is the transition operator for the resolvent chain and has the form

$$\mathcal{T}_\lambda = \int_0^\infty dr e^{-r} \Phi_{r,\lambda}.$$

There exist $L > 0$ and $0 < \epsilon < 1$ such that the following statements hold:

(i). For all $s, s' \in \Sigma$ with $H(s), H(s') \leq L$,

$$\mathcal{T}_\lambda(s, ds') > \epsilon \mu_L^{-1} ds'.$$

(ii). For all $s \in \Sigma$ with $H(s) > L$,

$$\int_{H(s') \leq L} \mathcal{T}_\lambda(s, ds') \geq \epsilon.$$

In the above, μ_L is defined to be the Lebesgue measure of the region $\{H(s) \leq L\} \subset \Sigma$, and we have inserted this constant for convenience. Given any $L > 0$, we can find an ϵ such that statement (i) is true by the argument in the proof of Part (1) of Proposition 4.3 (in which the cutoff was $l = 1 + \sup_x V(x)$ rather than an arbitrary fixed L , but this does not change the argument). The statements (i) and (ii) imply that the probability of jumping into the region $\{s \in \Sigma \mid H(s) \leq L\}$ is at least $\epsilon > 0$ from any starting point. The statements (i) and (ii) and the fact that $\Psi_{\infty, \lambda}$ is a stationary state for the dynamics are sufficient to prove the exponential ergodicity without further reference to the dynamics at hand. The proof of (ii) is below, and we proceed next with the remainder of the proof.

By (i), we have the minorization condition $\mathcal{T}_\lambda(s, ds') \geq h(s)\nu(ds')$ for

$$h(s) = \epsilon \chi(H(s) \leq L) \quad \text{and} \quad \nu(ds) = ds \mu_L^{-1} \chi(H(s) \leq L).$$

Let $\tilde{\sigma} \in \tilde{\Sigma} = \Sigma \times \{0, 1\}$ be the split chain defined using the pair h, ν above and \tilde{n}_m , $m \geq 1$ be the sequence of return-times to the atom. The times \tilde{n}_m form a delayed renewal chain in which the delay distribution is $\tilde{\mathbb{P}}_{\tilde{\Psi}}^{(\lambda)}[\tilde{n}_1 = n]$, and the jumps have distribution $\tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)}[\tilde{n}_1 = n]$.

We will treat the difference between $\Psi \mathcal{T}_\lambda^n$ and $\Psi_{\infty, \lambda}$ in the norm $\|\cdot\|_1$ through the formula (A.3). For $g \in B(\Sigma)$,

$$\begin{aligned} \int_{\Sigma} (\Psi \mathcal{T}_\lambda^n)(ds) g(s) &= \mathbb{E}_{\Psi}^{(\lambda)}[g(\sigma_n)] = \tilde{\mathbb{E}}_{\tilde{\Psi}}^{(\lambda)}[g(S_t)] \\ &= \tilde{\mathbb{E}}_{\tilde{\Psi}}^{(\lambda)} \left[\sum_{m=1}^n \sum_{r=1}^{\infty} \chi(m-1 = \tilde{n}_r) g(\sigma_n) \chi(\tilde{n}_{r+1} \geq n) \right] \\ &= \tilde{\mathbb{E}}_{\tilde{\Psi}}^{(\lambda)} [g(\sigma_n) \chi(\tilde{n}_1 \geq n)] + \sum_{m=1}^n F_{\tilde{\Psi}}^{(\lambda)}(m-1) \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)} [g(\sigma_{n-m}) \chi(\tilde{n}_1 \geq n-m)], \end{aligned}$$

where $F_{\tilde{\Psi}}^{(\lambda)} : \mathbb{N} \rightarrow \mathbb{R}^+$ is the renewal function

$$F_{\tilde{\Psi}}^{(\lambda)}(m) = \sum_{r=1}^{\infty} \tilde{\mathbb{P}}_{\tilde{\Psi}}^{(\lambda)}[m = \tilde{n}_r].$$

We will delay the demonstration of the following two statements until the end of the proof.

(I). There is a $c > 0$ such that for all $m \in \mathbb{N}$ and probability measures $\Psi \in M_+^1(\Sigma)$,

$$\tilde{\mathbb{P}}_{\tilde{\Psi}}^{(\lambda)}[\tilde{n}_1 \geq m] \leq ce^{-\epsilon m}.$$

(II). There is a $c > 0$ such that for all $m \in \mathbb{N}$ and probability measures $\Psi \in M_+^1(\Sigma)$,

$$\|F_{\tilde{\Psi}}^{(\lambda)}(m) - \Psi_{\infty, \lambda}(h)\|_{\infty} \leq ce^{-\epsilon m}.$$

With the above statements

$$\begin{aligned}
& \left| \int_{\Sigma} (\Psi \mathcal{T}_{\lambda}^n)(ds) g(s) - \Psi_{\infty, \lambda}(g) \right| \leq \left| \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_n) \chi(\tilde{n}_1 \geq n)] \right| \\
& + \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} |F_{\Psi}^{(\lambda)}(m-1)| \left| \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_{n-m}) \chi(\tilde{n}_1 \geq n-m)] \right| \\
& + \Psi_{\infty, \lambda}(h) \left| \sum_{m=\lfloor \frac{n}{2} \rfloor + 1}^n \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_{n-m}) \chi(\tilde{n}_1 \geq n-m)] - \frac{\Psi_{\infty, \lambda}(g)}{\Psi_{\infty, \lambda}(h)} \right| \\
& + \sum_{m=\lfloor \frac{n}{2} \rfloor + 1}^n |F_{\Psi}^{(\lambda)}(m) - \Psi_{\infty, \lambda}(h)| \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_{n-m}) \chi(\tilde{n}_1 \geq n-m)] \\
& \leq c \|g\|_{\infty} \left((1-\epsilon)^n + \epsilon^{-1} e^{-\epsilon \lfloor \frac{n}{2} \rfloor} + \epsilon^{-1} e^{-\epsilon \lfloor \frac{n}{2} \rfloor} \Psi_{\infty, \lambda}(h) + \epsilon^{-1} e^{-\epsilon \lfloor \frac{n}{2} \rfloor} \right), \tag{A.4}
\end{aligned}$$

where for the first inequality, we have applied the triangle inequality after adding and subtracting the term

$$\Psi_{\infty, \lambda}(h) \sum_{m=\lfloor \frac{n}{2} \rfloor + 1}^n \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_{n-m}) \chi(\tilde{n}_1 \geq n-m)].$$

The second equality in (A.4) uses (I) and (II), that $F_{\Psi}^{(\lambda)}(m) \leq 1$ for the second term, and the identity

$$\sum_{m=1}^{\infty} \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} [g(\sigma_m) \chi(\tilde{n}_1 \geq m)] = \tilde{\mathbb{E}}_{\Psi}^{(\lambda)} \left[\sum_{m=0}^{\tilde{n}_1} g(\sigma_m) \right] = \frac{\Psi_{\infty, \lambda}(g)}{\Psi_{\infty, \lambda}(h)}, \tag{A.5}$$

for the third term. The second equality of (A.5) follows by Part (2) of Proposition 4.14. The last line of (A.4) is $\|g\|_{\infty}$ multiplied by $O(e^{-\frac{n\epsilon}{2}})$, and so we have proven (A.3) for $\alpha = \frac{\epsilon}{2}$ when assuming statements (ii), (I), and (II).

(ii). We must return to a consideration of the process S_t underlying the resolvent chain. Let τ be a mean-1 exponential and $S_0 = s \in \Sigma$. For $H(s) > L$, let ς be the hitting time that $H(S_{\varsigma}) \leq L$. Since the Hamiltonian evolution preserves energy, we can give a lower bound for our quantity of interest as

$$\begin{aligned}
\int_{H(s') \leq L} \mathcal{T}_{\lambda}(s, ds') & \geq \mathbb{P}_s^{(\lambda)} [\varsigma \leq \tau \text{ and no collisions occur over the time interval } (\varsigma, \tau)] \\
& \geq \left(\inf_{H(x', p') \leq L} \frac{1}{1 + \mathcal{E}_{\lambda}(p')} \right) \mathbb{P}_s^{(\lambda)} [\varsigma \leq \tau] \tag{A.6}
\end{aligned}$$

The second inequality uses that the collisions occur with Poisson rate $\mathcal{E}_{\lambda}(S_t)$ and that $\tau - \varsigma$ is a mean-1 exponential when $\varsigma \leq \tau$ and conditioned on the information up to time ς . The escape rates $\mathcal{E}_{\lambda}(S_t)$ are uniformly bounded over any compact region, and it is thus sufficient for us to give a lower bound for $\mathbb{P}_s^{(\lambda)} [\varsigma \leq \tau]$. For any $T > 0$ that we choose,

$$\begin{aligned}
\inf_{H(s) > L} \mathbb{P}_s^{(\lambda)} [\varsigma \leq \tau] & = \inf_{H(s) > L} \int_0^{\infty} dt e^{-t} \mathbb{P}_s^{(\lambda)} [\varsigma \leq t] \\
& \geq e^{-T} - e^{-T} \sup_{H(s) > L} \mathbb{P}_s^{(\lambda)} [\varsigma > T].
\end{aligned}$$

If we show $\sup_{H(s) > L} \mathbb{P}_s^{(\lambda)}[\varsigma > T]$ is small for large T (or even merely bounded away from one), then combined with (A.6) the proof of statement (ii) is complete.

Define the function $w : \Sigma \rightarrow [\frac{1}{2}, 1]$ such that

$$w(s) = 1 - \frac{1}{2} \frac{1}{1 + H^{\frac{1}{2}}(s)}.$$

There exists L large enough so that for some $\delta > 0$

$$\sup_{H(x,p) > L} \int_{\mathbb{R}} dp' \mathcal{J}_\lambda(p, p') \left(w(x, p') - w(x, p) \right) \leq -\delta. \quad (\text{A.7})$$

Even though the rate of increase for $w(x, p)$ as $|p| \rightarrow \infty$ is gentle, it is still visible through the jump rates $\mathcal{J}_\lambda(p, p')$ in (A.7), because the escape rates $\mathcal{E}_\lambda(p) = \int_{\mathbb{R}} dp' \mathcal{J}_\lambda(p, p')$ grow on the order of $|p|$ by Part (1) of Proposition 3.1, and the landing p' for a jump from p with $|p| \gg 1$ is concentrated around $p' = \frac{1-\lambda}{1+\lambda}p$. In other words, when the momentum of the particle is high, a collision will occur quickly and the resulting momentum of the particle will typically be fraction $\frac{1-\lambda}{1+\lambda}$ of the former value.

The inequality (A.7) implies that $t\delta + w(S_t)$ behaves as a supermartingale when $H(S_t) > L$, since for $s = (X_t, P_t)$

$$\left. \frac{d}{dr} \right|_{r=0} \mathbb{E}_s^{(\lambda)}[r\delta + w(X_r, P_r)] = \delta + \int_{\mathbb{R}} dp' \mathcal{J}_\lambda(P_t, p') \left(w(X_t, p') - w(X_t, P_t) \right) \leq 0.$$

The equality follows since the function w is invariant of the Hamiltonian evolution. We have the following sequence of inequalities

$$\begin{aligned} \sup_{H(s) > L} \mathbb{P}_s^{(\lambda)}[\varsigma > T] &\leq \left(\sup_{H(s) > L} \mathbb{P}_s^{(\lambda)}[\varsigma > \delta^{-1}] \right)^{[\delta T]} \leq \left(\sup_{H(s) > L} \delta \mathbb{E}_s^{(\lambda)}[\varsigma \wedge \delta^{-1}] \right)^{[\delta T]} \\ &\leq \left(\sup_{H(s) > L} \mathbb{E}_s^{(\lambda)}[w(s) - w(S_{\varsigma \wedge \frac{1}{\delta}})] \right)^{[\delta T]} \leq 2^{-[\delta T]}. \end{aligned}$$

For the first inequality, the event $\varsigma > T$ requires that the time ς failed to occur over disjoint time intervals $(n\delta^{-1}, (n+1)\delta^{-1}]$ for $0 \leq n \leq [\delta T] - 1$. The second inequality is Chebyshev's and the third is by the optional stopping theorem, since $w(s) - w(S_t) - \delta t$ is submartingale over the time interval $t \in [0, \varsigma \wedge \delta^{-1}]$. For the last inequality, $\frac{1}{2} \leq w(s) \leq 1$. We can choose T to make the right side arbitrarily small.

(I). The second equality below follows from an inductive argument using the fact that the probability of the event $\tilde{\sigma}_n \in \Sigma \times 1$ is $h(\sigma_n)$ when conditioned on the information up to time $n-1$ and the value σ_n . This property is visible from the form of the transition operator $\tilde{\mathcal{T}}_\lambda$ given in Section 4.1.

$$\begin{aligned} \tilde{\mathbb{P}}_\Psi^{(\lambda)}[\tilde{n}_1 \geq m] &= \tilde{\mathbb{E}}_\Psi^{(\lambda)}[\chi(\tilde{n}_1 \neq 1) \cdots \chi(\tilde{n}_1 \neq m-1)] \\ &= \tilde{\mathbb{E}}_\Psi^{(\lambda)}[(1 - h(\sigma_0)) \cdots (1 - h(\sigma_{m-1}))] \\ &= \mathbb{E}_\Psi^{(\lambda)}[(1 - h(\sigma_0)) \cdots (1 - h(\sigma_{m-2}))] \mathbb{E}^{(\lambda)}[(1 - h(\sigma_{m-1})) \mid \sigma_0, \dots, \sigma_{m-2}] \\ &\leq (1 - \epsilon) \mathbb{E}_\Psi^{(\lambda)}[(1 - h(\sigma_0)) \cdots (1 - h(\sigma_{m-2}))]. \end{aligned}$$

The third equality identifies the expectation as from the original statistics, and the inequality uses that the probability σ_{m-1} jumps from σ_{m-2} to the support of h is $\geq (1 - \epsilon)$ by (i) and (ii). Applying this inductively, we get the bound.

(II). By the renewal theorem, $F_{\Psi}^{(\lambda)}(n)$ converges as $n \rightarrow \infty$ to the inverse of the expectation of the renewal increments $\tilde{n}_{m+1} - \tilde{n}_m$. The expectation of the increments $\tilde{n}_{m+1} - \tilde{n}_m$ is given by

$$\tilde{\mathbb{E}}_{\tilde{\Psi}}^{(\lambda)}[\tilde{n}_{m+1} - \tilde{n}_m] = 1 + \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}[\tilde{n}_1] = \tilde{\mathbb{E}}_{\tilde{\nu}}^{(\lambda)}\left[\sum_{n=0}^{\tilde{n}_1} 1\right] = \frac{1}{\Psi_{\infty, \lambda}(h)},$$

since the distribution for the split chain following a return to the atom is $\tilde{\nu}$ by Part (1) of Proposition 4.1, and where the second equality is by Part (1) of Proposition 4.4. The tails of the delay distribution $\tilde{\mathbb{P}}_{\tilde{\Psi}}^{(\lambda)}[\tilde{n}_1 > n]$ and the renewal jump distribution $\tilde{\mathbb{P}}_{\tilde{\nu}}^{(\lambda)}[\tilde{n}_1 > n]$ decay with order $O(e^{-\epsilon n})$ for large n by (I). It follows that the renewal function $F_{\Psi}(n)$ converges exponentially with rate ϵ to the value $\Psi_{\infty, \lambda}(h)$. □

References

- [1] C. S. Adams, M. Sigel, J. Mlynek: *Atom Optics*, Phys. Rep. **240**, 143-210 (1994).
- [2] K. B. Athreya, P. Ney: *A new approach to the limit theory of recurrent Markov chains*, Trans. Am. Math. Soc. **245**, 493-501 (1978).
- [3] A. Brunnschweiler: *A connection between the Boltzmann equation and the Ornstein-Uhlenbeck process*, Arch. Rational Mech. Anal. **76**, 247-263 (1981).
- [4] G. Birkel, M. Gatzke, I. H. Deutsch, S. L. Rolston, W. D. Phillips: *Bragg scattering from atoms in optical lattices*, Phys. Rev. Lett. **75**, 2823-2827 (1998).
- [5] K. L. Chung: *A course in probability theory*, Academic Press, 1976.
- [6] J. Clark: *Suppressed dispersion for a quantum particle in a δ -potential with random momentum kicks*, arXiv:1008.4502 (2010).
- [7] J. Clark: *A limit theorem to a time-fractional diffusion*, to appear.
- [8] J. Clark, L. Dubois: *Bounds for the state-modulated resolvent of a linear Boltzmann generator*, to appear.
- [9] J. Clark, C. Maes: *Diffusive behavior for randomly kicked Newtonian particles in a periodic medium*, Comm. Math. Phys. **301**, 229-283 (2011).
- [10] D. Dürr, S. Goldstein, J. L. Lebowitz: *A mechanical model for a Brownian motion*, Comm. Math. Phys. **78**, 507-530 (1981).
- [11] K. Dzharidze, E. Valkeila: *On the Hellinger type distances for filtered experiments*, Prob. Th. Rel. Fields **85**, 105-117 (1990).

- [12] M. I. Freidlin, A. D. Wentzell: *Random perturbations of Hamiltonian systems*, Mem. Amer. Math. Soc. **109**, no. 523, (1994).
- [13] N. Friedman, R. Ozeri, N. Davidson: *Quantum reflection of atoms from a periodic dipole potential*, J. Opt. Soc. Am. B **15**, 1749-1755 (1998).
- [14] P. Hall, C. C. Heyde: *Martingale limit theory and its application*, Academic Press (1980).
- [15] H. Hennion: *Sur le mouvement d'une particule lourde soumise à des collisions dans un système infini de particules légères*, Wahrsch. Verw. Geviete, 123-154 (1973).
- [16] R. Holley: *The motion of a heavy particle in a one dimensional gas of hard spheres*, Probab. Th. Rel. Fields **17**, 181-219 (1971).
- [17] R. Höpfner, E. Löcherbach: *Limit theorems for null recurrent Markov processes*, Mem. Amer. Math. Soc. **161**, (2003).
- [18] J. Jacod, A. N. Shiryaev: *Limit theorems for stochastic processes*, Springer Verlag Berlin, 1987.
- [19] S. Kunze, S. Dürr, G. Rempe: *Bragg scattering of slow atoms from a standing light wave*, Europhys. Lett. **34**, 343-348 (1996).
- [20] T. Komorowski, C. Landim and S. Olla: *Fluctuations in Markov Processes*. See webpage <http://w3.impa.br/~landim/notas.html>.
- [21] E. Löcherbach, D. Loukianova: *On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions*, Stoch. Proc. Appl. **118**, 1301-1321 (2008).
- [22] J. J. McClelland: *Atom-optical properties of a standing-wave light field*, J. Opt. Soc. Am. B **12**, 1761-1768 (1995).
- [23] S. P. Meyn, R. L. Tweedie: *Generalized resolvents and Harris recurrence of Markov processes*, Contemp. Math. **149**, 227-250 (1993).
- [24] E. W. Montroll, G. H. Weiss: *Random walks on lattices, II*, J. Math. Phys. **6** 167-181 (1965).
- [25] O. Morsch: *Dynamics of Bose-Einstein condensates in optical lattices*, Rev. Mod. Phys. **78**, 179-215 (2006).
- [26] E. Nelson: *Dynamical theories of Brownian motion*, Princeton University Press, 1967.
- [27] J. Neveu: *Potentiel Markovien récurrent des chaînes de Harris*. Ann. Inst. Fourier, Grenoble **22**, 7-130, 1972.
- [28] E. Nummelin: *A splitting technique for Harris recurrent Markov chains*, Z. Wahrscheinlichkeitstheorie Verw. Geb. **43**, 309-318 (1978).
- [29] D. Pollard: *Convergence of Stochastic Processes*, Springer-Verlag, 1984.
- [30] H. Spohn: *Large scale dynamics of interacting particles*, Springer-Verlag, 1991.

- [31] D. Szász, B. Tóth: *Towards a unified dynamical theory of the Brownian particle in an ideal gas*, **111** 41-62 (1987).
- [32] A. Touati: *Théorèmes limites pour les processus de Markov récurrents*, Unpublished paper (1988).
- [33] G. E. Uhlenbeck, L. S. Ornstein, *On the theory of Brownian motion*, Phys. Rev. **36**, 823-841 (1930).